PL-TR-96-2310

A COMPREHENSIVE APPROACH TO OUTLIER DETECTION AND EVENT CLASSIFICATION

H.L. Gray
S.R. Sain
W.A. Woodward

Southern Methodist University
Department of Statistical Science
Dallas, TX 75275-0235

December 1996

Scientific Report #3

PHILLIPS LABORATORY
DIRECTORATE OF GEOPHYSICS
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010

SPONSORED BY
Advanced Research Projects Agency (DoD)
Nuclear Monitoring Research Office
ARPA ORDER No. C-325

MONITORED BY
Phillips Laboratory
CONTRACT No. F19628-95-C-0098

This technical report has been reviewed and is approved for publication.

DELAINE R. REITER
Contract Manager
Earth Sciences Division

JAMES F. LEWKOWICZ
Director
Earth Sciences Division

This report has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

Qualifed requestors may obtain copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

If your address has changed, or you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify PL/IM, 29 Randolph Road, Hanscom AFB, MA 01731-3010. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1996	Scientific No.3

4. TITLE AND SUBTITLE

A Comprehensive Approach To Outlier Detection and Event Classification

5. FUNDING NUMBERS

PE 62301E
PR NM95
TA GM
WU AB
F19628-95-C-0098

6. AUTHOR(S)

H.L. Gray, S.R. Sain, and W.A. Woodward

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Southern Methodist University
Department of Statistical Science
Dallas, Texas 75275-0235

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Phillips Laboratory
29 Randolph Road
Hanscom AFB, MA 01731-3010
Contract Manager: Dr. Delaine Reiter/GPE

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

PL-TR-96-2310

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In this report, a comprehensive approach to outlier detection and event classification is investigated. Although the methodology is based on the assumption of available training data, it does not require ground truth or labels of any type. In fact, it is not even required that the number of different populations composing the training data is known. Data from Western China is analyzed to demonstrate the methodology, as well as some simulated data. These examples demonstrate vividly the importance of the role of correlation in selecting the best features. A method for feature selection is considered. Additionally, the problems of classifying events, numerical stability, missing data, signal to noise ratios, and mixture (discrete and continuous) data are discussed.

14. SUBJECT TERMS

Outlier detection, classification, feature selection, singularities, missing date, mixture distributions.

15. NUMBER OF PAGES

60

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

CONTENTS

A Comprehensive Approach to Outlier Detection and Classfication

H.L. Gray, Stephan R. Sain and Wayne A. Woodward

Department of Statistical Science, Southern Methodist University

1. Introduction and Background

In previous technical reports we have considered the problem of automated outlier detection under a number of different scenarios. Initially, we considered the problem of detecting an outlier from a given population (say earthquakes or explosions) when a training set of labeled data (i.e. its source is known) was available. These results were successfully demonstrated by Baek, Gray, McCartor, and Woodward (1992), Fisk, Gray, and McCartor (1993, 1994) and Fisk and Gray (1993). In a later report Wang, Woodward, Gray, Wiechecki, and Sain (1996) have extended this methodology to the case in which the training data could be a mixture of any number of event types. In that report, it was, however, assumed that the number of different types of events was known and that at least some of the events were labeled. That is, it was assumed that at least some ground truth was available for the training data. Sain, Gray, and Woodward (1996) have considered the case in which the training data can again be composed of any number of event types, but no labels are known, i.e. no ground truth is required. Additionally, they drop the additional assumption that the number of different types of events is known. This scenario is considered further here.

In this report a more comprehensive approach is considered which not only addresses the outlier problem but considers:

a. selecting the best features

b. classifying/assigning labels

c. removing singularities to eliminate numerical problems

Additional capabilities will be added later to this comprehensive package including adjusting for missing data and differences in signal-to-noise ratio, and the use of discrete feature variables. These items are discussed briefly in the section on future research.

2. Technical Discussion

Our previous results are based on the assumption of the existence of a training sample, $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$, from a population ($\pi_1$) of interest (e.g. earthquakes or a mixture of earthquakes and non-nuclear or industrial explosions) along with an additional (and possibly suspicious) observation, $\boldsymbol{U}$, to be tested as an outlier from π_1. That is, the hypothesis

$$\begin{aligned} &\mathrm{H}_0: \boldsymbol{U} \in \pi_1 \\ &\mathrm{H}_1: \boldsymbol{U} \notin \pi_1 \ . \end{aligned} \tag{1}$$

is tested using a generalized likelihood (GLR) ratio test. In the work of Wang, et al. (1996) and Sain, et al. (1996), we assume that the training data consists of a sample of size n from a mixture distribution whose density is given by

$$f(x) = \sum_{i=1}^{m} p_i g_i(\boldsymbol{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \tag{2}$$

where m is the number of components in the mixture, $g_i(\boldsymbol{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ is the density associated with the ith component, the p_i, $i = 1, \ldots, m$, are the mixing proportions, and $\boldsymbol{x}$ is a d-dimensional vector of feature variables. As mentioned above, a typical scenario might be the case in which the mixture population consists of events associated with

earthquakes and mining explosions. The hypotheses in (1) is tested using a modified likelihood ratio statistic, W, given by

$$W = \frac{\sup_{\theta \in \Theta} L_0(\theta)}{\sup_{\theta \in \Theta} \widetilde{L}_1(\theta)} \tag{3}$$

where

$$L_0(\theta) = \left(\prod_{s=1}^{n} f(\boldsymbol{X}_s; \boldsymbol{\theta}) \right) f(\boldsymbol{U}; \boldsymbol{\theta}),$$

and

$$\widetilde{L}_1(\theta) = \prod_{s=1}^{n} f(\boldsymbol{X}_s; \boldsymbol{\theta}).$$

Since the null distribution of W has no known closed form, the bootstrap is used to approximate critical values. It should also be noted that, by making use of the bootstrap technique, this procedure requires no distributional assumptions concerning the outlier distribution. This is a useful practical solution because of the lack of regional training samples for nuclear events. The rejection region is thus of the form $W \leq W_\alpha$ where W_α is picked to provide a level α test. The maximum likelihood estimates are obtained using the expectation-maximization (EM) algorithm (see Little and Rubin, 1987), and initial estimates are required for this iterative algorithm. The results of Wang, et al. (1996) are based on a known number of components (i.e. m in (2) is known) and enough labeled training data in order to provide these starting values.

In the setting of interest, it is probable that little none of the training data will be labeled and that even the number of non-nuclear event types in the region may be unknown. Current results by Sain, Gray, and Woodward (1996) show that a modification

of the Wang et al. (1996) procedure to this setting of interest is feasible. When none of the training data are labeled, they suggest a clustering approach to group the data into distinct classes from which initial estimates can be obtained. When the number of event types is unknown, Sain et al. (1996) consider the use of AIC (Akaike, 1974) for purposes of determining the number of components m in the mixture. Specifically, for $m = 1, \ldots, M$, the AIC criterion is defined by

$$\text{AIC}(m) = -2\ln(L_{max}(m)) + 2(\#\text{ of free parameters}) \tag{4}$$

where $L_{max}(m)$ is the maximized likelihood of the training sample under the assumption that there are m components, M is a sufficiently large integer, and the last term is a penalty imposed to avoid overfitting. Parameter estimates are obtained via the EM algorithm. For each m, $m = 1, \ldots, M$, a hierarchical clustering routine is used to obtain m initial groups to provide starting values for the EM algorithm. The AIC criterion is calculated for $m = 1, \ldots, M$, and the number of components, m_{AIC}, associated with the minimum AIC value is chosen. The test statistic for the data, W, is then calculated based on m_{AIC} components. The distribution of W in this setting is obtained using the bootstrap. The authors also examined the use of BIC (Akaike, 1977) which imposes a more stringent penalty and has better asymptotic behavior than AIC. The findings based on simulations were that in the cases considered, the use of either AIC or BIC combined with initial clustering produced results that are surprisingly comparable to those based on knowing m and the availability of some labeled data. Additionally, our simulation results seem to favor AIC over BIC although both methods tend to produce satisfactory results. For specific results see Sain, et al. (1996).

a. Feature Selection

In most settings there will eventually be a large number of available feature variables with which to perform the analyses. We have examined the problem of identifying the "best" features for use. Since the goal is to obtain features that best separate the outlier population from the population of the training data, it is not surprising that the selection of "best" features requires some knowledge of both populations. Our initial solutions to the feature selection problem have used a forward stepwise procedure that begins by finding the "best" single feature. The best single feature is defined to be that feature for which the power of detecting the outlier in a given set of alternatives is the largest for fixed false alarm rate. Once the best single feature is obtained, a second feature is added that maximizes this power for two features. Features are added until a a predetermined maximum number of features is attained or until the resulting power begins to decrease. It is important to note that in many instances, use of all available features is not the optimal choice in terms of maximizing power. It is also important to understand that the best set of k features will not necessarily consist of the best k features considered separately. In this regard, the role of correlation among the feature variables is emphasized by considering the simulated example involving training samples with hypothetical features as shown in Figure 1. Here we see a case in which the features are independent (left-hand side) and a case in which they are highly correlated (right-hand side). In the top row, a potential outlier indicated by the solid dot at (-1, 1) is not determined to be an outlier in the case of independent features (upper left-hand) while it would be classified as an outlier in the highly correlated setting despite not really being an unusual observation in either of the univariate dimensions taken separately. Thus, in this case, if the outlier population of interest were centered around (-1, 1) then the two highly correlated features would be the preferred set of features since points around (-1, 1) are highly unlikely to occur in their bivariate distribution as can be seen in the figures. The very opposite type of behavior is shown in the plots on the bottom row. That is, the

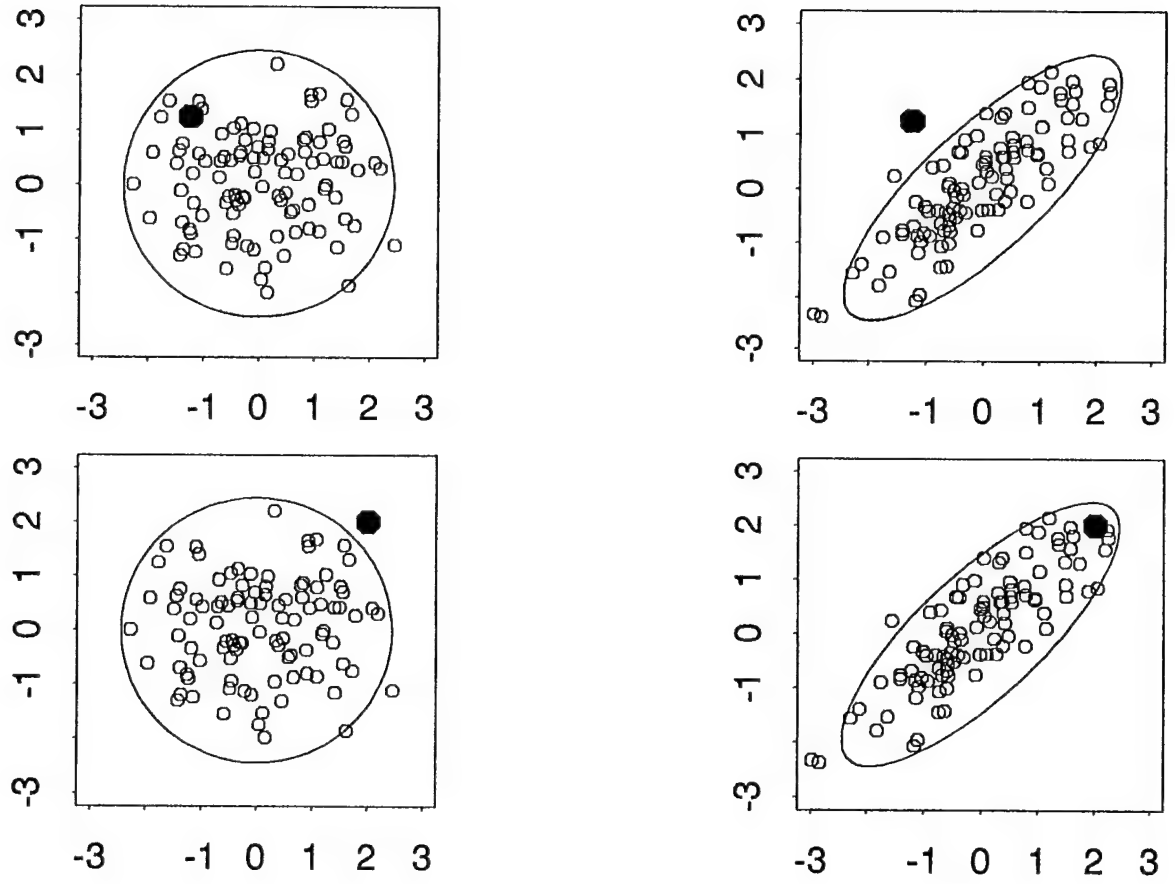

Figure 1

potential outlier at (2, 2) would be classified as an outlier in the case of independent variables but is not as unusual when considering the distribution of highly correlated variables. The lesson to be learned from such an example is that correlation as well as the position of the outlier must be taken into account when considering variable selection.

What is really needed is a robust procedure for variable selection that requires little or no information about the outlier population, and this will be a focus of our research in this area. These remarks will be demonstrated further when we consider some real data in Section 3.

b. Classifying/Assigning Labels

It was noted that when the training sample consists of a single (non-mixture) population, the outlier test automatically classifies an observation as being a member of the training sample population (e.g earthquakes) or not. However, it is clear that when the training sample comes from a mixture distribution, failing to reject the null hypothesis does not indicate which component of the mixture to which the event should be assigned. The same problem actually arises with the training sample observations themselves in this mixture setting. Simply knowing that observation $\boldsymbol{X}_j$ belongs to the training sample and is non-nuclear does not indicate what type of event it is. We have performed investigations into the labeling of training sample events and new events determined to not be outliers, and developed methodology for this purpose when there are two or more populations to be labeled.

We first consider the case in which the training data consist of two components where the goal is to label the observations in the training sample. Each point in the training sample is tested as an outlier from each of the two training sample components and corresponding p-values obtained are associated with each component. Based on these p-values, each training sample member would be assigned a component membership or will be left unassigned when membership is not clear as defined by some

predetermined p-value. Whereas most multivariate tests are inherently nondirectional, we have developed and are continuing to develop tests based on a focused critical region (see Schucany, Frawley, Wang, and Gray, 1996). This report is included in the appendix. Such tests can be used to improve our ability to assign component membership based on the position of the training sample value being tested with respect to the locations of the corresponding component centroids. When the distribution of the training sample has more than two components, the testing can be based on considering the components two at a time. Actual "naming" of components can be done by an analyst, or by a defined statistic and/or auxiliary variables.

c. Removing Singularities to Eliminate Numerical Problems

When d features measured on n events are detected at J stations, then when dJ is large, data compression may be required. Let X_{jki} denote the measurement of the kth feature for the ith event in the training sample measured at the jth station. That is, for the kth feature, we have the following training data:

Station 1	...	**Station J**
X_{1k1}	...	X_{Jk1}
$\vdots$		$\vdots$
X_{1kn}	...	X_{Jkn}

We use the notation $\overline{X}_{jk}$ to denote the average of the n events measured at station j and feature k and $\overline{\boldsymbol{X}}_{\boldsymbol{k}} = (\overline{X}_{1k}, \ldots, \overline{X}_{Jk})'$ to denote the vector of these averages evaluated at each of the J stations. The J station readings for the potential outlier at the kth variable are denoted by $\boldsymbol{U_k} = (U_{1k}, \ldots, U_{Jk})'$.

In the non-mixture setting, Fisk, Gray, and McCartor (1995) and Gray, Woodward, and Yücel (1995) considered several strategies for dealing with multi-station data. When Jd is not too large, an obvious approach would be the "full-vector" approach in which the d features at each of the J stations are considered as a single vector consisting of Jd feature variables, i.e.

$$\boldsymbol{X_i} = (X_{11i}, X_{12i}, \ldots, X_{1di}, X_{21i}, X_{22i}, \ldots, X_{2di}, \ldots, X_{J1i}, X_{J2i}, \ldots, X_{Jdi})' . \quad (5)$$

A new observation to be tested as an outlier is then a similarly configured $Jd \times 1$ vector denoted $\boldsymbol{U}$. This full-vector approach is not a data compression approach, and when Jd is large, this approach may not be feasible. Another approach considered by Fisk, et al. (1995) and Gray, et al. (1995) included declaring an event to be an outlier if any of the individual station-based tests finds it to be an outlier using a Bonferroni-based adjustment to assure that the overall significance level is no larger than α. They also investigated methods of compressing the data by calculating new "features", that are linear combinations of the observations on feature k at each of the J stations. The outlier detection is then based on a likelihood ratio test as before but is calculated using only the d new variables. Fisk et al. (1995) and Gray, et al. (1995) considered weights chosen to minimize the variance of the resulting feature. Simulation studies by these authors found that the power of the full-vector approach was the best of the procedures considered since it was more robust and consistently competitive with the other procedures. However, due to the potential dimensionality problems with the full-vector approach, data-compression alternatives have continued to be explored. Recent results by Woodward, Wang, Gray, and Frawley (1996) provide compression weights that give results that are similar to those of the full-vector approach. The weights used for the kth feature are designed so that the distance between the mean of the compressed training data and the compressed potential outlier is a maximum. They also considered a two-stage compression

approach. Their initial simulation results indicate that these new compression procedures perform comparably with the full vector approach and are thus to be preferred when dimensionality is a problem for the full-vector approach.

Wang, et al. (1996) have developed a GLR outlier test for the more realistic case in which there may be more than one type of non-nuclear event in a region, and the training sample is from this mixture distribution. Their work allows for the training sample to represent a sample from, for example, earthquakes and mining blasts. If data are collected on d features at J stations, then if Jd is not large, as in the non-mixture setting, the full-vector approach would be a viable approach. However, dimensionality may prove to be a problem and alternative approaches have been considered. As in the non-mixture case, an individual station-based approach would be possible that declares an event to be an outlier if any of the individual station-based tests (using the Wang et al. (1996) GLR test) finds the event to be an outlier using a Bonferroni-based adjustment. Recall that in the non-mixture case this procedure proved to be inferior to the full-vector approach. Unfortunately, in the mixture case, we have been unable to find weights analogous to those considered by Woodward, et al. (1996) for optimally combining stations to produce maximum separation between the outlier population and the population of the training sample. However, preliminary results show that a viable approach to data compression in the mixture setting is as follows. For a given station j, one can consider the readings from the d features as a d-dimensional "feature" vector on which the modified likelihood ratio test of Wang et al. (1996) can be applied to obtain $W(j)$. In this case, large values of $W^{-1}(j)$ are suggestive of an outlier. Letting $\boldsymbol{H} = (W^{-1}(1), \ldots, W^{-1}(J))'$, we calculate $D_H = \boldsymbol{H}'\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{H}}^{-1}\boldsymbol{H}$ as an overall measure of the size of the $W^{-1}(j)'s$. Thus, large values of D_H suggest that the observed value is an outlier. We can then use a bootstrap approach to approximate its null distribution. It should be noted that the original sample of size n produces only a single observation on $\boldsymbol{H}$, and because of this we use a separate bootstrap step to calculate $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{H}}$.

Specifically, we obtain B_1 nonparametric bootstrap samples of size $n+1$ from the original training sample, and from each of these samples we calculate $\boldsymbol{H}$. We then let $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{H}}$ be the sample variance/covariance matrix. We then take B_2 nonparametric bootstrap samples of size $n+1$ from the original training sample in order to find the null distribution of D_H using the $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{H}}$ obtained from the first bootstrap step. Specifically, for purposes of the hypothesis test, the $100(1-\alpha)$th percentile of $D_H^*(b)$, $b = 1, \ldots, B_2$ is found. Ideally, given a bootstrap sample for which D_H is to be calculated, a second bootstrap sample would be taken from this sample in order to obtain a bootstrap-based estimate of $\boldsymbol{\Sigma}_{\boldsymbol{H}}$ specific to that sample. However, this procedure would be very computationally intensive, and we have thus chosen the faster method of simply calculating $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{H}}$ once and using this estimate for each of the B_2 bootstrap samples.

Preliminary simulation results have shown that this procedure has merit. We have considered a hypothetical case of two features and two stations. For station j, $j = 1, 2$, the distribution of the training sample follows the mixture model

$$f_j(\boldsymbol{x}) = 0.5 g_{1j}(\boldsymbol{x}; \boldsymbol{\mu}_{1j}, \boldsymbol{\Sigma}_j) + 0.5 g_{2j}(\boldsymbol{x}; \boldsymbol{\mu}_{2j}, \boldsymbol{\Sigma}_j) \tag{6}$$

where $g_{rj}(\boldsymbol{x}; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_j)$ is the (multivariate normal) density associated with the rth component, and $\boldsymbol{x}$ is a 2-dimensional vector of feature variables. Note that the two components each have the same covariance. In the simulations, $\boldsymbol{\mu}_{1j} = (-1, 1)$ and $\boldsymbol{\mu}_{2j} = (1, -1)$ for each station. The outlier populations were assumed to be MVN with variance/covariance matrix I and with the mean vector of the outlier population k given by $(k-4, k-4)$, $k = 1, \ldots, 4$. In Table 1 we consider the case in which $\Sigma_1 = \Sigma_2 = I$ and the case $\Sigma_1 = I$ and $\Sigma_2 = 4I$ where I is the identity matrix, where in both cases considered here the stations are assumed to be independent. In Table 1 we compare the full-vector results with those based on D_H. There it can be seen that the test based on D_H had power comparable to that using the full-vector test.

It should be noted that these results are very preliminary, and we are continuing to study data compression in the mixture setting in the presence of differing correlation structures between stations and feature variables.

Table 1. Power Comparison Between D_H and Full Vector Test

	k				
	1	2	3	4	
D_H	.998	.901	.299	.037	$\Sigma_1 = \Sigma_2 = I$
Full Vector	1.000	.893	.362	.075	
D_H	.974	.668	.179	.043	$\Sigma_1 = I,\ \Sigma_2 = 4I$
Full Vector	.966	.706	.234	.070	

3. Example Using Seismic Data

In order to clarify the remarks in the previous section we consider a data set furnished by Dr. Steve Taylor of LANL. The data consists of log(Pg/Lg) ratios at various frequency bands obtained from events at WMQ in western China. The training data consist primarily of earthquakes along with what may be a few commercial explosions. Additionally, 18 nuclear explosions from the region are also available.

For clarity of exposition, our initial analysis (presented in the current and next two paragraphs) uses only two features. Following this we consider the more general problem which includes selection of the best features. The two features considered initially are log(Pg/Lg) ratios for bands 1-2Hz and 4-8Hz. Feature selection will be considered later and is used to choose the "best" set of feature variables. For these data, AIC and BIC picked three and one populations respectively as the most likely mixture. This is illustrated in Figure 2. In order to examine these fits, we generated parametric bootstrap samples from the 3-component and 1-component models picked by AIC and BIC

respectively. The number of components chosen by AIC and BIC were obtained for each of the samples from the 1-component model, and both procedures tended to correctly find one component. However, for the samples from the 3-component model, AIC tended to pick three components while BIC (incorrectly) continued to pick a 1-component model for the majority of these samples. Furthermore, resampling from the original data (nonparametric bootstrap) resulted in AIC usually picking three components while BIC picked one. These comments are illustrated in Figure 3. These and other simulations suggest that AIC's estimate of three components is the more reliable estimate of the number of components although both fits were satisfactory.

In Figure 4, we show the scatterplots of the two features for the training data along with the components of the AIC and BIC mixture distribution fits and associated contours of the mixture distributions. While it is unclear whether the three components selected by AIC have a physical interpretation, it is clear that the mixture method provides a flexible model that will allow for considerable non-normality without increasing the false-alarm rate.

In Figure 5, we show the scatterplot of training data along with the corresponding points for the nuclear explosions. There it can be seen that for these data, there is strong separation between the training sample and the "outliers" to be tested. It is not surprising that the test based on W correctly detected all of these points as outliers (using either AIC or BIC). In order to examine the power of the outlier test when the "outlier population" is not as widely separated from the training data as the one for the current data, we moved the outlier population closer to the center of the training sample along the path indicated in the figure. The letters $(a - k)$ indicate the positions of the outlier population means considered. The sample covariance from the nuclear data was used to generate samples at each of these locations. A simulation study was performed to study the power of the

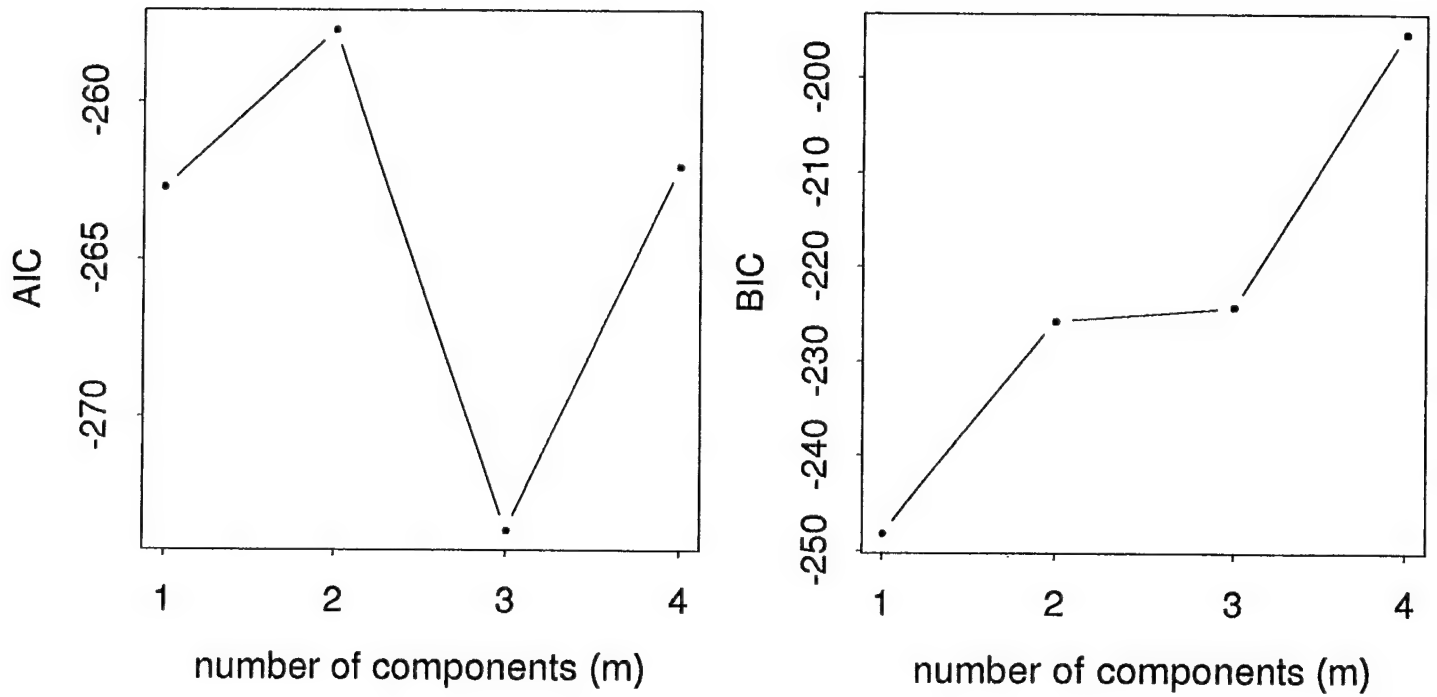
AIC
-260
-265
-270
1
2
3
4
number of components (m)
BIC
-200
-210
-220
-230
-240
-250
1
2
3
4
number of components (m)

Figure 2

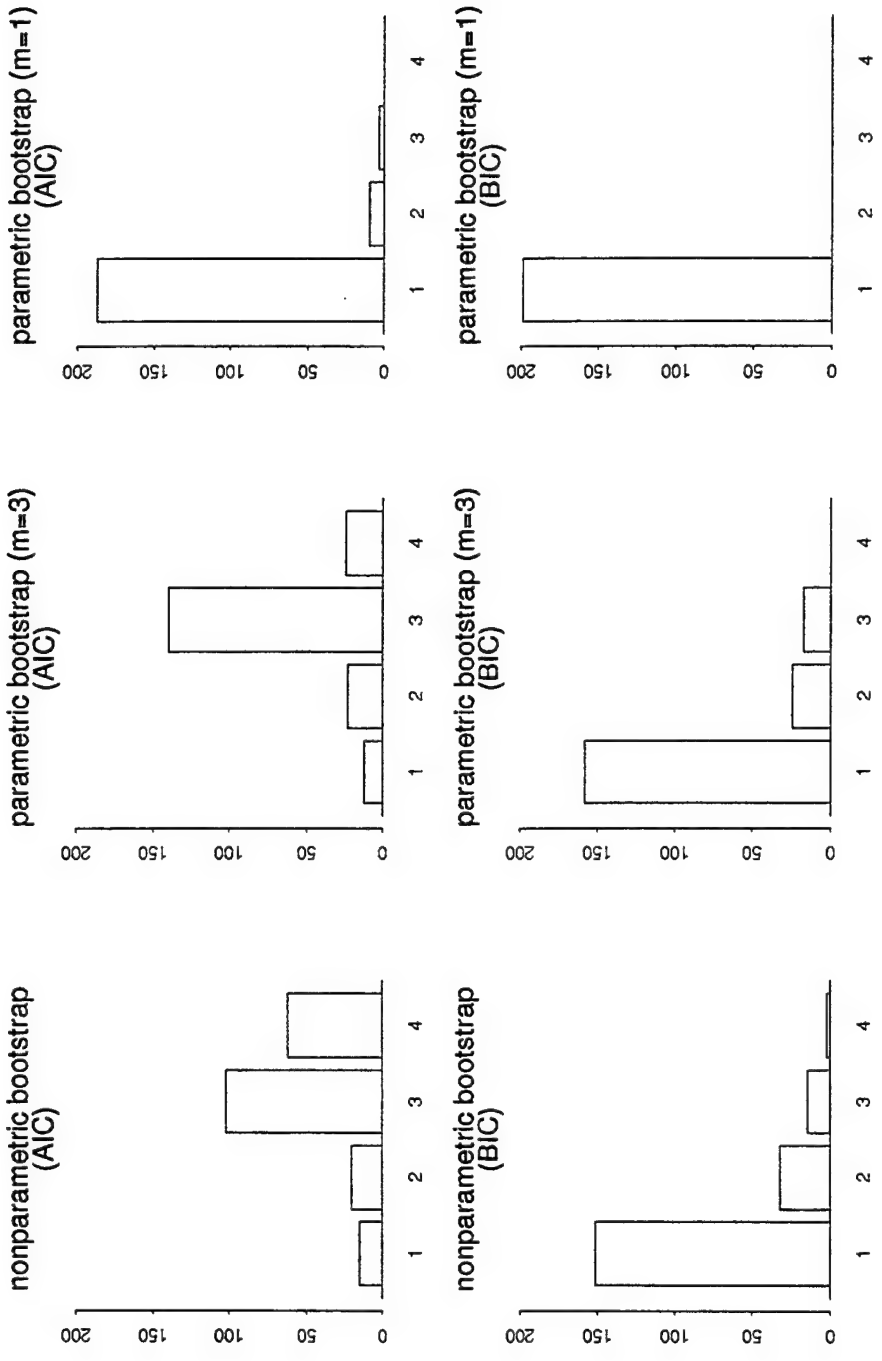
parametric bootstrap (m=1) (AIC)
parametric bootstrap (m=1) (BIC)
parametric bootstrap (m=3) (AIC)
parametric bootstrap (m=3) (BIC)
nonparametric bootstrap (AIC)
nonparametric bootstrap (BIC)
0
50
100
150
200
1
2
3
4

Figure 3

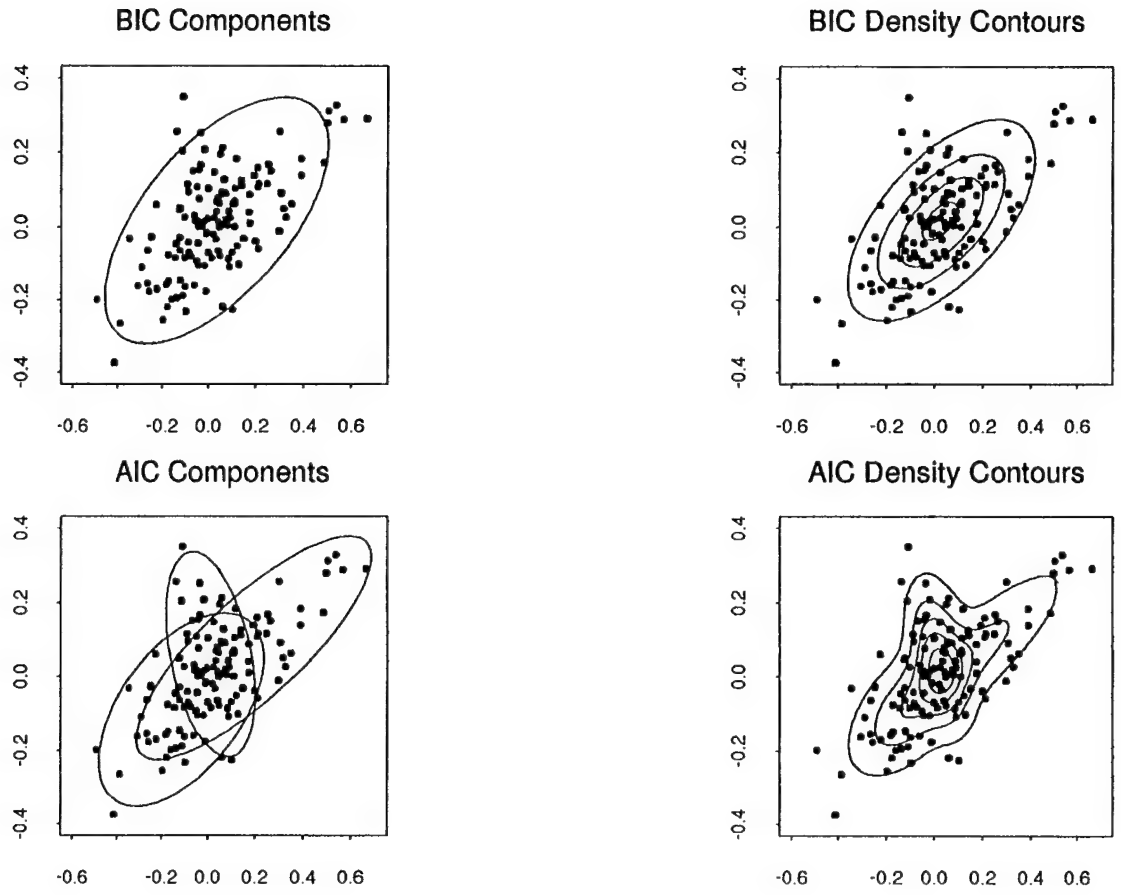
BIC Components
BIC Density Contours
AIC Components
AIC Density Contours
-0.6
-0.2
0.0
0.2
0.4
0.6
-0.4
-0.2
0.0
0.2
0.4

Figure 4

outlier test (using AIC and BIC) for outlier points from each of the artificial outlier populations. For widely separated outliers ($h - k$), both procedures were always able to properly detect them. Also, when the outlier population was too close to the training data ($a - c$) the outliers were rarely detected, as is to be expected. For outliers between these extremes, ($e - g$), the use of the three component model for the training data obtained by AIC resulted in much larger powers. These results are displayed in Figure 6. It should be noted that these results are dependent on the position (and in particular the direction) of the outlier population.

The WMQ data set included log(Pg/Lg) ratios for seven different frequency bands. Using the stepwise feature selection procedure discussed previously in Section 2, three features were chosen. These were the log(Pg/Lg) ratios for bands 0.5-1, 1.5-3, and 4-8Hz. It is interesting to note that these three features were chosen as best based on either the AIC or BIC choice of number of mixture components, and that in both cases the resulting outlier detection power decreased with the use of additional features. Also of interest is the fact that even though the high frequency ratios tended to best separate the populations when considered individually, the best set of three variables is composed of a low, middle, and high frequency ratio, and the best two features are a low and high frequency ratio. A correlation analysis and the position of the outlier population shows why the low and middle frequency ratios are included in the set of best features. In Figure 7, univariate densities are shown for log (Pg/Lg) ratios in the 0.5-1 Hz, 3-6 Hz, and 4-8 Hz frequency bands. In these plots, earthquakes are denoted with solid lines and nuclear events with dashed lines. Note that the low frequency shows no separation between the two event types while the two high frequencies show considerable separation. In Figure 8, we show the scatterplots for the training data (denoted "Q") and nuclear data (denoted "X") for two sets of feature variables. In Figure 8(a), we consider the log (Pg/Lg) ratios in frequency bands 0.5-1 Hz and 4-8 Hz while in Figure 8(b) we show a similar plot using log(Pg/Lg) ratios in the two separate high frequency bands, 3-6

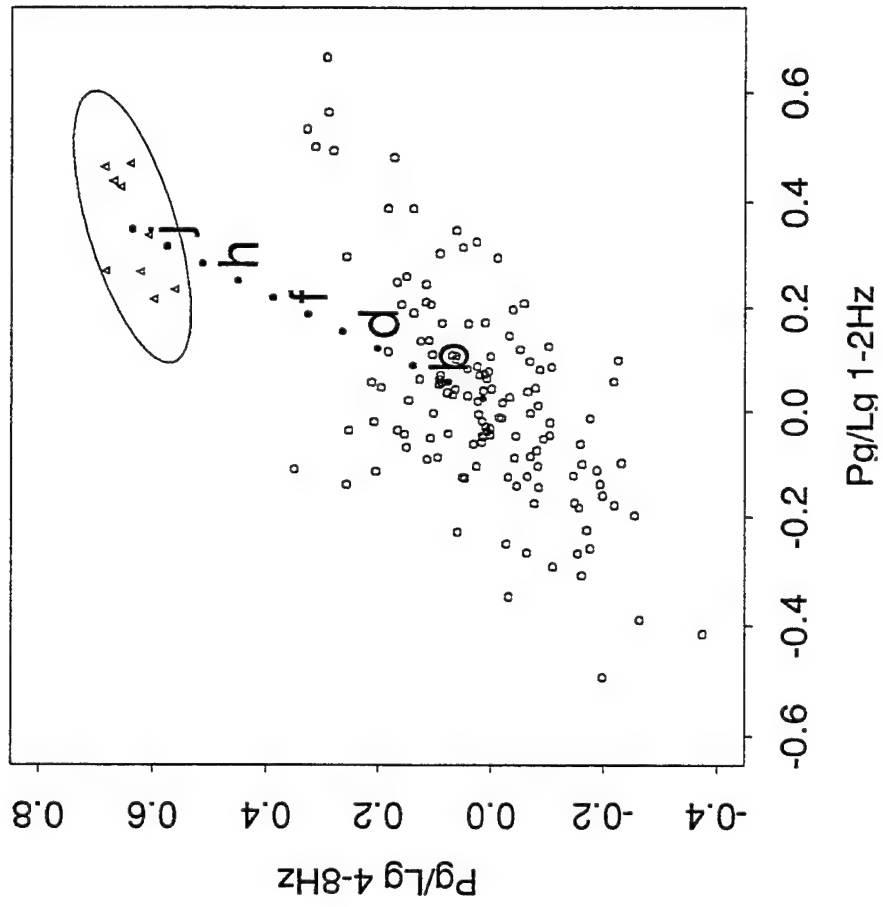

Figure 5

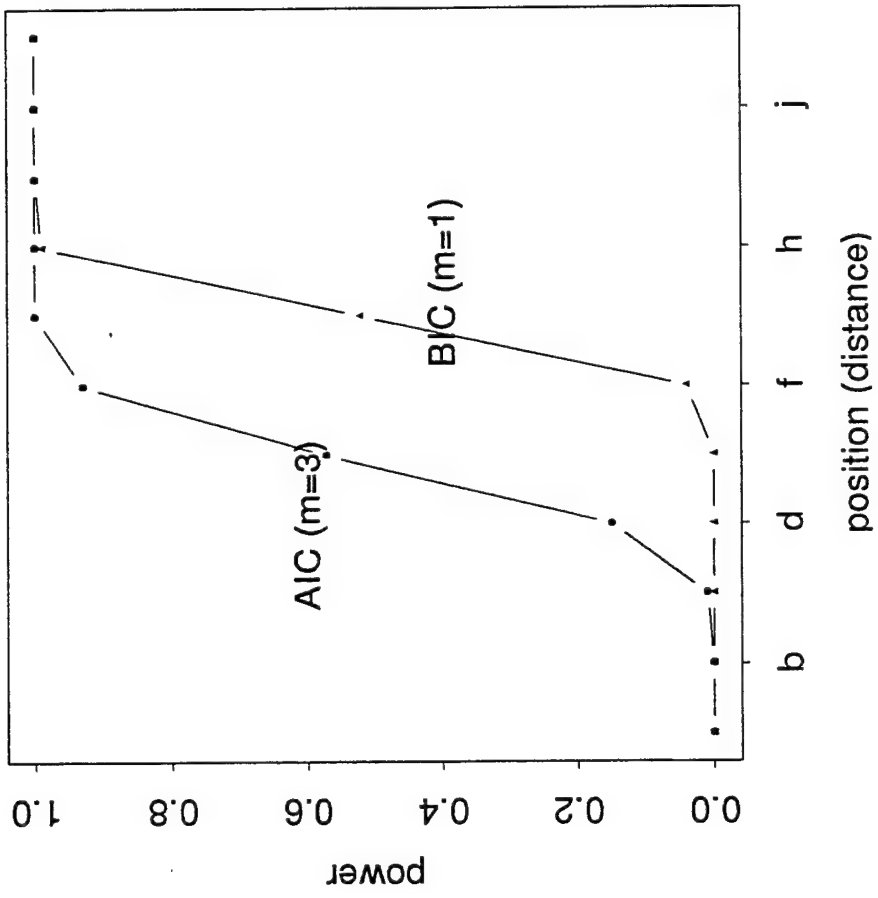

Figure 6

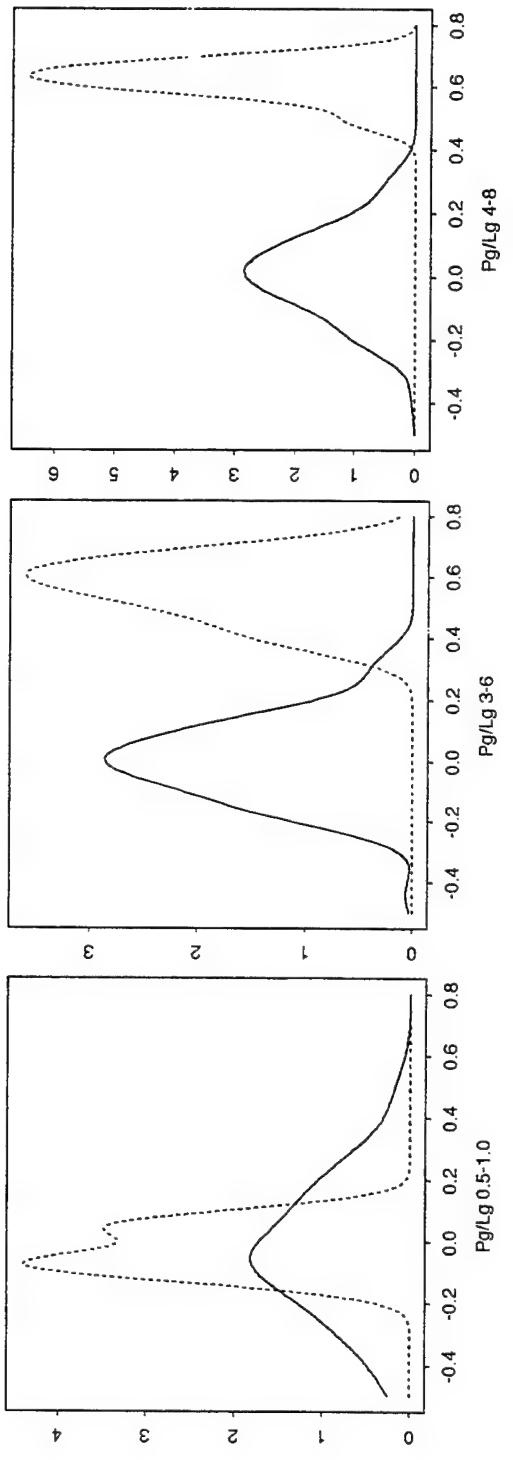

Figure 7

Hz and 4-8 Hz. In Figure 8(a), it can be seen that although the low-frequency ratio is a poor discriminator individually, its use in conjunction with a high frequency ratio enhances the degree to which the nuclear data is separated from the training data better than using the high frequency variable alone. In Figure 8(a) it can be seen that the nuclear data is separated from and not consistent with the correlation structure present in the training data. In Figure 8(b), it is seen that the two high frequency ratios are providing highly correlated, redundant information. While the separation between the training data and the nuclear data is still apparent, this separation was essentially no greater than that which would have been observed using a single high-frequency ratio, and hence the two high frequencies when considered together did not perform as well as the low and high frequencies considered jointly.

In this example, although seven features were available, a particular subset of three was selected as best. The methodology used for selection can be used when the "direction of interest" of the outlier population is known for the features of interest. However, this may not be the case and more robust methods will need to be developed.

In order to demonstrate the labeling (or classification) to the WMQ data we artificially add a second component that is somewhat closer to the training data but in the general direction of the nuclear data. This would approximate the case of two major components in the training data (e.g. earthquakes and explosions). In Figure 9, scatterplots of the previous training data along with the artificial second component are shown. In our analysis we treat this expanded data set as the new training sample. Not surprisingly AIC and BIC picked four and two components respectively. The components and contours for the fits are also shown in Figure 9. For simplicity, using $m = 2$ as selected by BIC we then applied the labeling procedure using a focused critical region test to classify the training sample elements. The results of this classification are shown in Figure 10 where the component in which the points were placed is indicated with a "1" or "2" while unclassified points are indicated with "0". It should be noted that

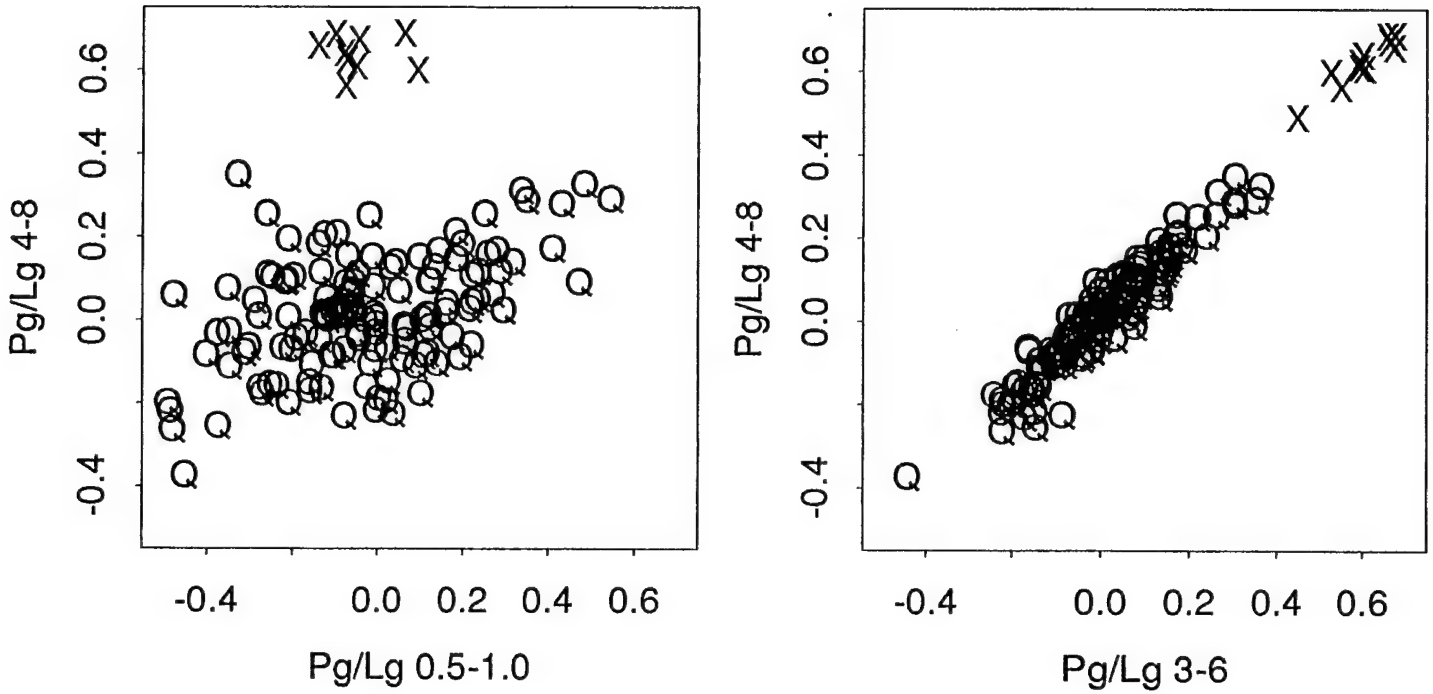
Pg/Lg 4-8
-0.4
0.0
0.2
0.4
0.6
-0.4
0.0
0.2
0.4
0.6
Pg/Lg 0.5-1.0
Pg/Lg 4-8
-0.4
0.0
0.2
0.4
0.6
-0.4
0.0
0.2
0.4
0.6
Pg/Lg 3-6

Figure 8

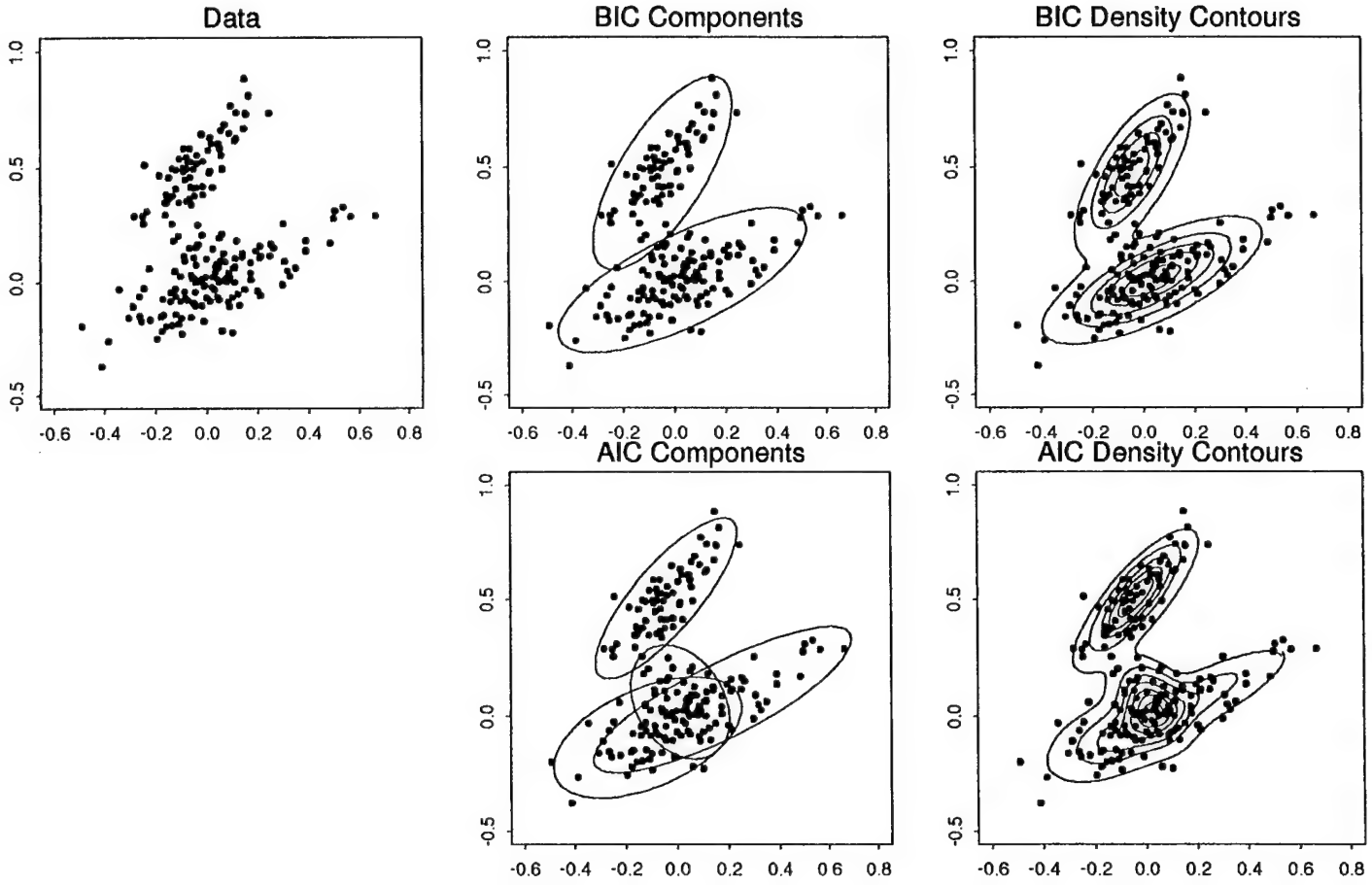
Data
BIC Components
BIC Density Contours
AIC Components
AIC Density Contours
1.0
0.5
0.0
-0.5
-0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8

Figure 9

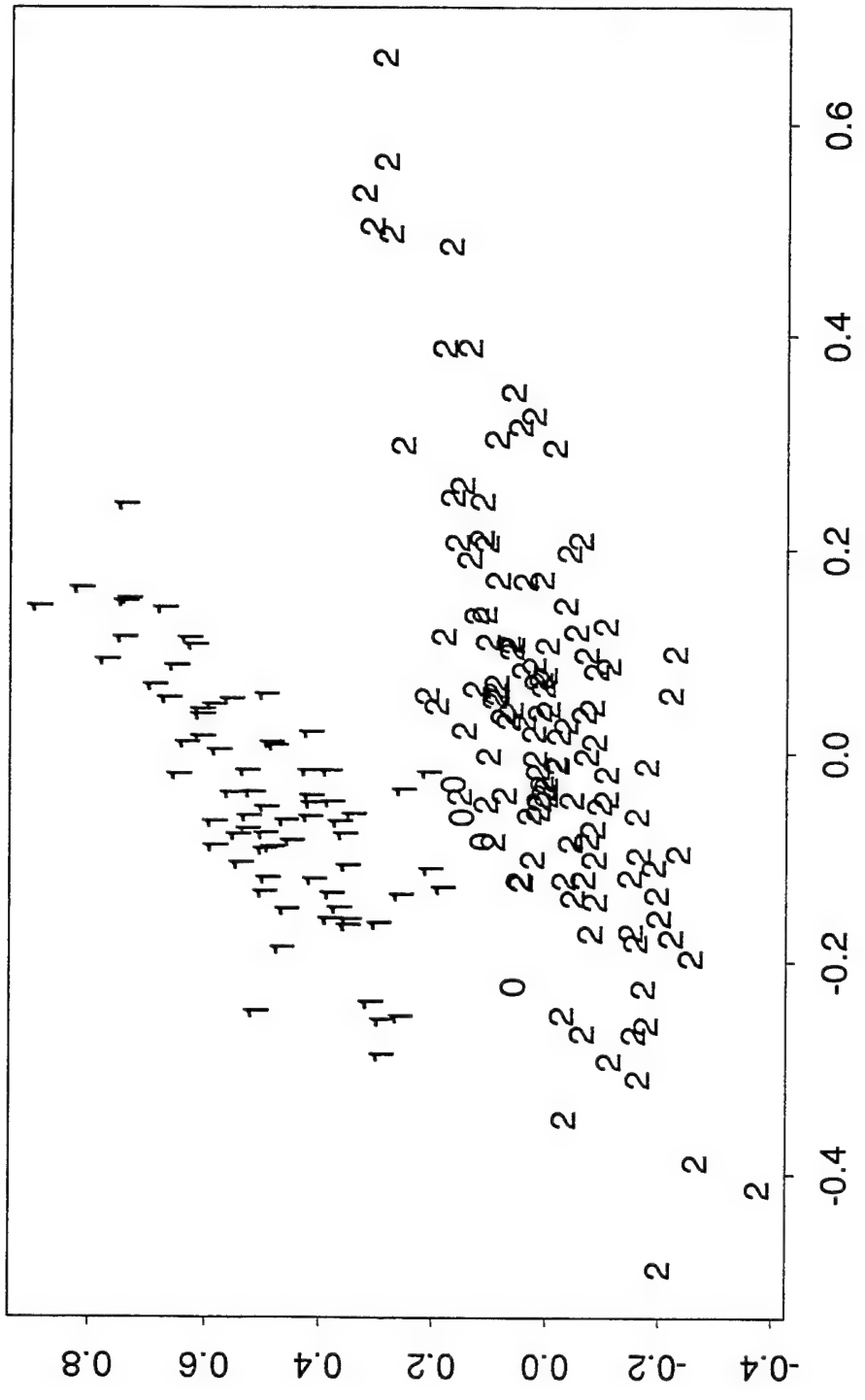

Figure 10

these unclassified events are still in the mixture (i.e. not outliers) but there is insufficient evidence to classify them into one group or the other. It can easily be seen that the classification has done a good job of classifying the points into their correct groups, which are known in this case.

4. Future Research

Although much has been accomplished, there is much yet to be done in order for the methodology developed and being developed to function properly and robustly in the expected environments. Several issues which we will be addressing in the near future include modifying the GLR method for mixtures to adjust for:

(a) signal-to-noise ratio

(b) missing data

(c) use of both continuous and discrete features

Each of these items is briefly discussed in the remaining parts of this report.

(a) Signal-to-Noise Ratio

In all of our previous results, although the variance/covariance across realizations was properly accounted for in the GLR method, the signal-to-noise ratios for the waveforms were assumed to be adequate and comparable. This may not be the case, and therefore it will be necessary to adjust the various features for signal-to-noise ratio before feeding the data to the GLR outlier and classification programs.

As an example suppose the observed feature $\widetilde{X}_i$ is

$$\widetilde{X}_i = \frac{\mathrm{RMS}(Pg_i+n_i)}{\mathrm{RMS}(Lg_i+n_i)} \tag{7}$$

for Pg_i and Lg_i in some given frequency band where n_i is the associated noise. Then, if the noise is uncorrelated with the signal, we have $\mathrm{Var}(Pg+n) = \mathrm{Var}(Pg) + \mathrm{Var}(n)$ and $\mathrm{Var}(Lg+n) = \mathrm{Var}(Lg) + \mathrm{Var}(n)$. Thus, $\mathrm{Var}(Pg) = \mathrm{Var}(Pg+n) - \mathrm{Var}(n)$ and $\mathrm{Var}(Lg) = \mathrm{Var}(Lg+n) - \mathrm{Var}(n)$, so that

$$\left(\frac{\mathrm{Var}(Pg)}{\mathrm{Var}(Lg)}\right)^{1/2} = \left(\frac{\mathrm{Var}(Pg+n)-\mathrm{Var}(n)}{\mathrm{Var}(Lg+n)-\mathrm{Var}(n)}\right)^{1/2}. \tag{8}$$

But (8) suggests the corrected feature

$$X_i = \left(\frac{\mathrm{MS}(Pg_i+n_i)-\mathrm{MS}(n_i)}{\mathrm{MS}(Lg_i+n_i)-\mathrm{MS}(n_i)}\right)^{1/2}, \tag{9}$$

and as we see, this feature variable is adjusted for signal/noise ratio and estimates the true feature ratio, shown on the left side of (8). The feature X_i could now be used in the GLR method in place of $\widetilde{X}_i$. The same argument could be made for any feature based on the ratio of RMS values and in that case the noise variables need not be the same (as would be the case in cross-spectral ratios). That is, if $\widetilde{X}_i$ is given by

$$\widetilde{X}_i = \frac{\mathrm{RMS}_1(f_{1i}+n_i)}{\mathrm{RMS}_2(f_{1i}+a_i)}, \tag{10}$$

where $\sigma_{n_i}^2 \neq \sigma_{a_i}^2$, then a similarly determined X_i would be used in the tests.

These approaches and others will be investigated so that the signal-to-noise ratio of each feature will be properly incorporated into all of the GLR outlier routines and classification methods.

(b) Missing Data

Actual seismic data will often involve missing observations. We have previously investigated the use of the EM algorithm (see Little and Rubin, 1987) for dealing with missing data in the case of outlier detection where the training sample comes from a single (non-mixture) population and a single station is used (see Miller, et al., 1993, 1994). Interestingly, one of the findings of Miller, et al. (1993, 1994) was that the use of a simple mean-replacement procedure performed as well as the use of the full EM algorithm in this case. Of course, if the full vector approach to multiple stations is used, then the previously developed missing data analysis will apply. However, when data must be compressed, appropriate methods of handling missing data need to be developed.

Additionally, techniques for appropriately dealing with missing data in the case of multiple populations in the training sample need to be developed. Preliminary results by Sain, et al. (1996) indicate that the use of mean replacement after a clustering based on available data has the effect of artificially reducing the estimated component variances which results in inflated significance levels. Thus, it appears that the use of the EM algorithm will required in this case to improve performance, and we propose to investigate this application. If we consider the mixture density as in (2) where $\boldsymbol{X}$ is a d-component feature vector, then actually this mixture model can be considered to involve a $d+1$st categorical variable that defines component membership. In the case in which some or all of the training data are unlabeled, this categorical variable has missing observations. The EM algorithm used by Wang, et al. (1996) and Sain, et al. (1996) accomplishes the appropriate maximization in the presence of the missing data concerning membership. Our future research will examine the use of the EM algorithm to handle not only the missing component membership information but also missing data on the feature variables.

(*c*) Discrete and Continuous Feature Variables

The use of both discrete and continuous feature variables were included in the results by Baek, et al. (1992) and Gray, et al. (1996) for the problems of outlier detection and of classification when the training sample originated from a single (non-mixture) population. Additionally, Miller, et al. (1994) considered this situation when some data were missing.

Unfortunately, those results do not apply directly to the more general mixture case which was introduced to eliminate the need for ground truth in outlier detection. The use of both discrete and continuous feature variables in the multiple population (mixture) setting will be addressed in the upcoming research in order to also eliminate the need for ground truth in the case of mixed continuous and discrete data. Similarly, the use of discrete and continuous feature variables in data compression will also be addressed both for the single population and multiple population settings.

References

Akaike, H. (1974), "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control* 19, 716-723.

Akaike, H. (1977), "On Entropy Maximization Principle," *Applications of Statistics*, Ed. P.R. Krishnaiah, Amsterdam: North-Holland, 27-41.

Baek, J., Gray, H.L., McCartor, G.D., and Woodward, W.A. (1992). "A Generalized Likelihood Ratio Test in Outlier Detection or Script Matching," DARPA Semi-Annual Technical Report.

Fisk, M. D. and Gray, H.L.(1993). "Event Identification Analysis of the Novaya Zemlya Event on 31 December 1992 using Outlier and Classification Likelihood Ratio Tests," (Paper in volume compiled by Ryall, 1993), MRC-R-1449, Mission Research Corp., Santa Barbara, CA.

Fisk, M.D., Gray, H.L. and McCartor, G.D. (1993). "Applications of Generalized Likelihood Ratio Tests to Seismic Event Identification," PL-TR-93-2221, Phillips Laboratory, Hanscom AFB, MA, ADA279479.

Fisk, M.D., Gray, H.L. and McCartor, G.D. (1994). "Preliminary Assessment of Seismic CTBT/NPT Monitoring Capability," Proceedings of the 16th Annual Seismic Research Symposium, September 1994, PL-TR-94-2217, ADA284667.

Fisk, M., Gray, H.L., and McCartor, G.D. (1995), "Statistical Methodology and Assessment of Seismic Event Characterization Capability," Mission Research Corporation, Phillips Lab PL-TR-95-2156, ADA305487.

Gray, H.L., Woodward, W.A., and Yücel, Z.T. (1995), "Outliers with Multiple Stations," Southern Methodist University Technical Report, Department of Statistical Science, February, 1995.

Little, R.J.A. and Rubin, D.B. (1987), *Statistical Analysis with Missing Data,* New York: John Wiley and Sons.

Miller, J.W., Gray, H.L., and Woodward, W.A. (1993), "Discriminant Analysis and Outlier Testing when Data are Missing," Semi-Annual Technical Report Four, Advanced Research Projects Agency, Nuclear Monitoring Research Office, August 19, 1993.

Miller, J.W., Gray, W.A., and Woodward, W.A. (1994), "A Hypothesis Testing Approach to Discriminant Analysis with Mixed Categorical and Continuous Variables when Data are Missing," SMU-TR-273, July 1994.

Sain, S.R., Gray, H.L., and Woodward, W.A. (1996), "Outlier Detection without Ground Truth," *Proceedings of the18th Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty,* PL-TR-96-2153, ADA313692..

Sain, S.R., Gray, H.L., and Woodward, W.A. (1996), "Outlier Detection when Training Data are Unlabeled," in Phillips Laboratory technical report PL-TR-96-2049, ADA308972.

Schucany, W.R., Frawley, W.H., Wang, S. and Gray, H.L. (1996), "Multivariate Testing with Positive Orthant Alternatives," submitted to *Journal of the American Statistical Association.*

Wang, S., Woodward, W.A., Gray, H.L., Wiechecki, S., and Sain, S.R. (1996), "A New Test for Outlier Detection from a Multivariate Mixture Distribution," *Journal of Computational and Graphical Statistics*, accepted to appear.

Woodward, W.A., Wang, S., Gray, H.L., and Frawley, W.H. (1996), "A New Look at Outlier Tests with Multiple Stations," in Phillips Laboratory technical report PL-TR-96-2049, ADA308972.

APPENDIX

"Multivariate Testing with Positive Orthant Alternatives" by Schucany, Frawley, Wang, and Gray. Submitted to *Journal of the American Statistical Association.*

Multivariate Testing with Positive Orthant Alternatives

William. R. Schucany, William. H. Frawley, Suojin Wang, and H. L. Gray,

ABSTRACT

The likelihood-ratio test (LRT) of the null hypothesis that a multivariate mean equals zero versus the positive-orthant alternative is reexamined. Perlman (1969) derived the exact null distributions under normality for general cone alternatives. However, because these distributions depend on the unknown covariance matrix, the usable critical points have only been bounds. For important cases the resulting one-sided tests are biased. The disappointing performances of these approximate LRT have been the subject of several critical articles over the years. We show that the bootstrap can rescue the LRT by estimating the appropriate critical point. Monte Carlo comparisons confirm its superiority to Hotelling's T^2, a "half-space" alternative investigated by Tang (1994) and closely related simple test due to Follmann (1996). The proposed nonparametric test should perform well when the distribution is not multivariate normal. In addition, it is easy to extend the methodology to general cone-shaped critical regions.

Key Words: Bootstrap, Cone, Likelihood-ratio, Non-normality, One-sided

1. INTRODUCTION

First consider the special case where the $\mathbf{x}_i$, $i = 1,...,n$ are a p-dimensional random sample from a normal population with unknown mean μ and covariance matrix Σ. Suppose that one is testing the null hypothesis that the mean of the population is located at zero versus the alternative that at least one of the elements of the mean is non-null. It is well known

---

William R. Schucany is Professor of Statistical Science, William H. Frawley is Director of the Center for Statistical Consulting, and H. L. Gray is Frensley Professor of Mathematical Sciences, all at Southern Methodist University, Dallas, Texas 75275. Suojin Wang is Associate Professor of Statistics at Texas A&M University, College Station, Texas 77843. This research was sponsored in part by ARPA Contract No. F19628-95-C-0098.

that the likelihood-ratio test (LRT) in this instance leads to the Hotelling T^2 statistic and, subsequently, an F statistic with p and n - p degrees of freedom. We will address the nonnormal case with our proposed approach.

There are instances in which one may be interested only in detecting specific departures of the mean vector from the null hypothesis. For example, if the data are responses of subjects to a stimulus in an experiment, then the effects may only be of interest if the individual elements of the mean are located in a specific direction from the origin. Without loss of generality, the direction of interest can be selected to be the positive orthant, where at least one of the μ_i is positive.

Again under normality Perlman (1969) developed the theory for a LRT with one-sided alternatives (cones). His LRT has been the subject of many articles since it appeared. The exact null distribution depends on the covariance matrix. Consequently, the usable critical values have only been bounds. For important classes of alternatives, such as for the positive orthant case, these bounds produce biased tests. A recent article to address this difficulty is by Tang (1994), who shows that these same critical points are exact for a "half space" alternative. For a variety of values of p and n, Tang empirically compared the improvement in power of the LRT, with a particular half space alternative, over the T^2 test.

This same one-sided alternative is examined by Follmann (1996). The new simple test introduced there uses traditional approximations to Hotelling's T^2 along with a requirement that the sum of the components in the sample mean vector be positive. The effect is that observed significance levels (either exact or approximate) are divided by 2. While this new test does have greater sensitivity than the classical omnibus test, it is quite specifically for the multivariate normal distribution. The new test is certainly simple, but somewhat ad hoc in focusing upon the sum of the components. The parametric counterpart that $\Sigma\mu_j > \mathbf{0}$ contains not only the positive orthant, which is of interest to us, but also more of the parameter space that is not. This is essentially Tang's halfspace without the more difficult requirement of maximizing constrained likelihoods and Perlman's exact critical points. More importantly, all of these rely heavily upon

multivariate normality for their validity, as do the more technical recent results of Berk and Marcus (1996).

The primary difficulty associated with applying Perlman's LRT when the cone is the positive orthant is the lack of an exact critical point under the null distribution. The bootstrap method (Efron (1979), Fisher and Hall (1990)) provides a solution to this difficulty, at least for samples of "moderate" size (e.g., n 10 for $p = 2$). A second difficulty is algorithmic in nature, in that as the dimension increases, computation of the LRT statistic becomes more complex. High-speed processing enables the use of a straightforward approach to this obstacle.

In this paper we apply the LRT of Perlman to the positive orthant cone, *without assuming normality*, by employing nonparametric bootstrapping to estimate the critical point under the null hypothesis. Details of calculating the test statistic are given in Section 3. Monte Carlo results are presented in Section 4 to illustrate the gains in power relative to the T^2 and Tang's half space alternative. Even though the normality assumption holds for our simulations, the bootstrap is superior without using that information.

2. THE LIKELIHOOD-RATIO TEST WITH A HALF SPACE ALTERNATIVE

A *half space*, H_u^+, is a set of the form $\{\mathbf{v} \mid \mathbf{v}'\mathbf{u} \quad 0\}$ for some fixed vector $\mathbf{u}$. A set is *one-sided* if it is contained in the interior of a half space. A *cone* is a positively homogeneous, closed, and one-sided set. The problem is to test the null hypothesis that $\boldsymbol{\mu} = \mathbf{0}$ versus the alternative hypothesis that $\boldsymbol{\mu} \in C$, where C is a cone.

Perlman derived the likelihood ratio statistic to be

$$U(\mathbf{x},\mathbf{m},A,C) = \|\mathbf{m}\|_A^2 \, (1 + \|\mathbf{m} - \mathbf{x}\|_A^2)^{-1}, \qquad (1)$$

where $\mathbf{x}$ denotes $n^{1/2}$ times the sample mean, A is $(n - 1)$ times the sample covariance matrix, and $\mathbf{m}$ $(= \widehat{\boldsymbol{\mu}}_c)$ is the vector in C that is closest to $\mathbf{x}$ in terms of the Mahalanobis distance

$$\|\mathbf{m} - \mathbf{x}\|_A^2 = (\mathbf{x} - \mathbf{m})' A^{-1}(\mathbf{x} - \mathbf{m}). \qquad (2)$$

Assuming normality Tang (1994) used a result of Perlman (1969) to calculate critical points for the LRT statistic $U(\mathbf{x},\mathbf{m},A,H_u^+)$, and compared, via Monte Carlo results, the power of $U(\mathbf{x},\mathbf{m},A,H_J^+)$ versus the Hotelling T^2. Here H_J^+ is the half-space defined by the vector $\mathbf{J} = (1,1,...,1)'$.

These quadratic forms can be motivated from far more general considerations than normal likelihood. As an early example, Fisher (1938) described his linear discriminant as the result of maximizing the ratio of two quadratic forms, which represent the variation between and within groups. This is true no matter what types of variables are elements of the vectors, including combinations of discrete and continuous. For some evidence and discussion of these points see Titterington, et al (1981). Consequently, our extension has much broader application than to the multivariate normal setting. Nevertheless, it is interesting to see how helpful it is in the special case.

3. EXTENSION TO THE POSITIVE ORTHANT ALTERNATIVE HYPOTHESIS

It is difficult (perhaps impossible) to analytically calculate critical points for the test statistic $U(\mathbf{x},\mathbf{m},A,Q^+)$ in which C is the positive orthant Q^+. However, bootstrapping provides one means of estimating the critical points of the statistic under the null hypothesis that the mean is $\mathbf{0}$. The process for conducting a test of size α, using nonparametric bootstrapping, is described in the following steps.

Procedure for the Bootstrap Test of the Null Hypothesis that the Mean is Zero Versus the Alternative that the Mean is in the Positive Orthant

1. Compute the $U(\mathbf{x},\mathbf{m},A,Q^+)$ statistic using the original set of observations.

2. Subtract the mean of the sample from each of the observations, leaving the n residuals.

3. Draw with replacement a random sample of size n from the residuals (for

3. Draw with replacement a random sample of size n from the residuals (for which the null hypothesis holds conditionally).

4. For that sample, compute the statistic $U(\mathbf{x}^*,\mathbf{m}^*,A^*,Q^+)$, where the * denotes the same quantities in (1) evaluated on the (re)sampled residuals.

5. Repeat steps 3 and 4 a total of B times and save the values of U^*.

6. Calculate the $(1 - \alpha)$th quantile of these B realizations, $u_{1-\alpha}$, the bootstrap estimate of the critical value.

7. If the statistic of step 1 equals or exceeds $u_{1-\alpha}$, then the null hypothesis is rejected in favor of the alternative that $\boldsymbol{\mu} \in Q^+$.

The main difficulty in carrying out the preceding steps is the calculation of the statistic in (1). If all the elements of the vector $\mathbf{x}$ are non-negative, then $\mathbf{m} = \mathbf{x}$, the denominator of (1) is unity, and the statistic is essentially the standard T^2 statistic. If, however, some or all of the elements of $\mathbf{x}$ are negative, then it remains to find the vector $\mathbf{m}$ in Q^+ which is closest to $\mathbf{x}$ in terms of the Mahalanobis distance. To aid in understanding the process, the methodology for the two-dimensional case will first be examined and then a general procedure will be described.

When $p = 2$, let the μ_1 axis be the abscissa and the μ_2 axis be the ordinate. The positive quadrant, where values of $\mu_1$and μ_2 are positive, is labeled as Quadrant 1. As usual, Quadrants 2 - 4 are found by moving counter-clockwise. If the first element of $\mathbf{x}$ is positive and the second element is negative, then $\mathbf{x}$ is located in Quadrant 4. This example is depicted in Figure 1, which shows $\mathbf{x}$ at the center of an ellipse which is touching the μ_1 axis at the point $\mathbf{m}_1 = (m_1,0)$. The shape of the ellipse is governed by the elements of A^{-1}. The vector $\mathbf{m}_1$ is the closest point to $\mathbf{x}$ on the μ_1 axis in the sense that it minimizes the distance

$$q(\mathbf{m}_1) = (\mathbf{x} - (m_1,0)')'A^{-1}(\mathbf{x} - (m_1,0)'). \qquad (3)$$

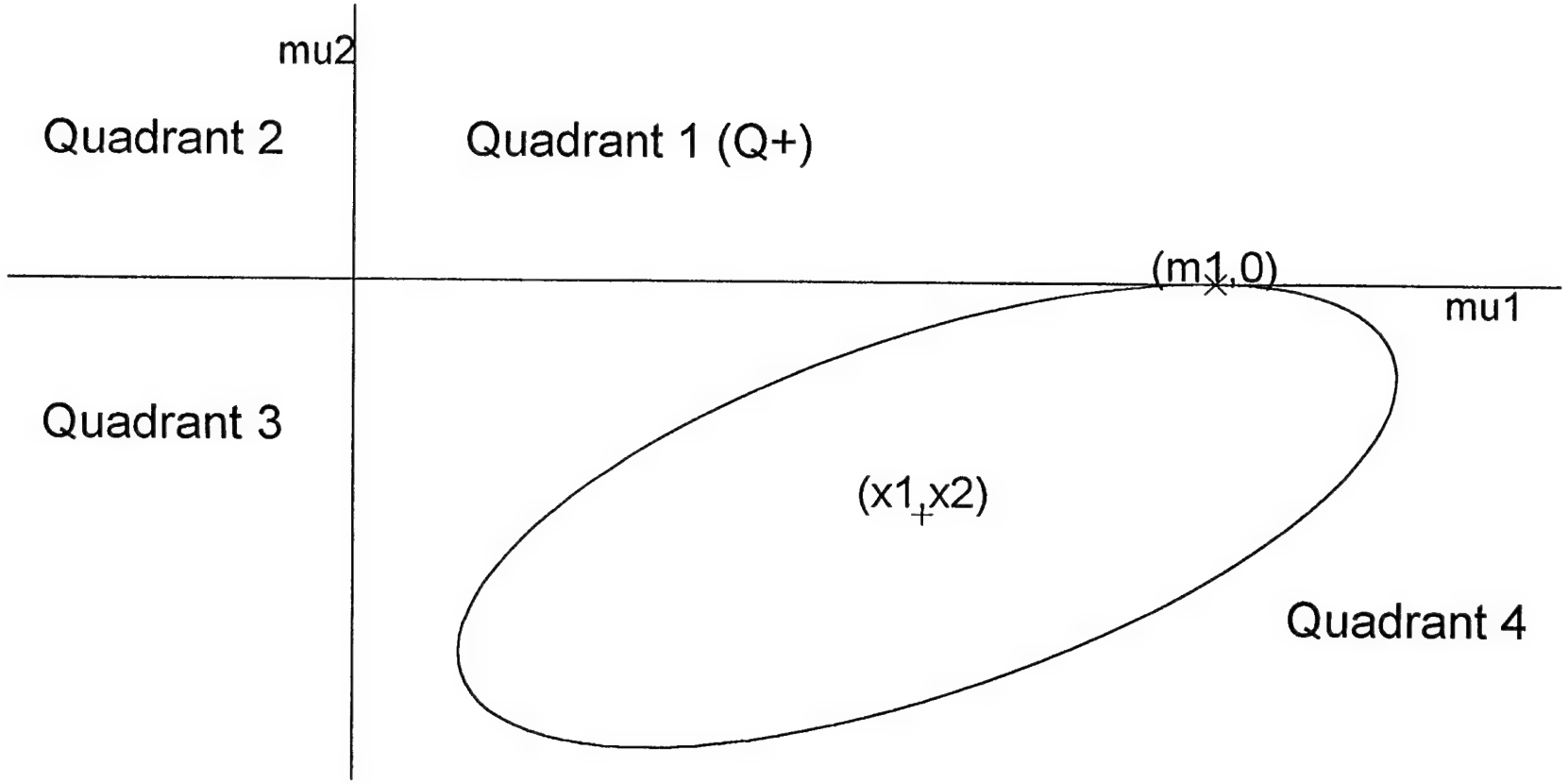

Figure 1. Closest Point in Q+ to Point in Quadrant 4

If the elements of the A^{-1} matrix are denoted by c_{ij}, $i,j = 1,2$, then the value of m_1 which minimizes (3) can be found by elementary calculus to be

$$m_1 = (c_{12}x_2 + c_{11}x_1)/c_{11}. \tag{4}$$

In the figure, $(m_1,0)$ is located on the non-negative part of the μ_1 axis, and thus is a legitimate candidate for the closest point in Q^+ to $\mathbf{x}$. In general, if m_1 were found to be negative, then the point (0,0) would be the point in Q^+ closest to $\mathbf{x}$ and on the μ_1 axis .

There is also a point $\mathbf{m}_2 = (0,m_2)$ that is the closest point to $\mathbf{x}$ on the μ_2 axis. It is the point which minimizes the distance

$$q(\mathbf{m}_2) = (\mathbf{x} - (0,m_2)')'A^{-1}(\mathbf{x} - (0,m_2)'), \tag{5}$$

and again by calculus

$$m_2 = (c_{12}x_1 + c_{22}x_2)/c_{22}. \tag{6}$$

If m_2 were found to be negative (which it would be in Figure 1), then the point (0,0) would be the point in Q^+ closest to $\mathbf{x}$ and on the μ_2 axis. The value of $\mathbf{m}$ that is used in (1) is the one associated with the smaller of $q(\mathbf{m}_1)$ and $q(\mathbf{m}_2)$; otherwise, $\mathbf{m} = \mathbf{0}$, if neither $\mathbf{m}_1$ nor $\mathbf{m}_2$ are in Q^+.

In summary, if $\mathbf{x}$ does not fall in the first quadrant, then depending on the quadrant in which it does fall and its sample covariance structure, there may not be a point $\mathbf{m}$ on the border of Q^+ other than the point (0,0), in which case $U(\mathbf{x},\mathbf{0},A,Q^+)$ is equal to zero. The algorithmic approach for calculating the statistic $U(\mathbf{x},\mathbf{m},A,Q^+)$ in two dimensions is summarized in the following steps:

Algorithm for Calculating $U(\mathbf{x},\mathbf{m},A,Q^+)$ in Two Dimensions

1. Set $\mathbf{m} = \mathbf{0}$ and $d_{\min} = \|\mathbf{x}\|_A^2$.
2. Calculate m_1 from (4). If $m_1 < 0$, go to step 4.
3. Calculate $q(\mathbf{m}_1)$. Set $\mathbf{m} = (m_1,0)'$ and $d_{\min} = q(\mathbf{m}_1)$.
4. Calculate m_2 from (6). If $m_2 < 0$, go to step 6.
5. Calculate $q(\mathbf{m}_2)$. If $q(\mathbf{m}_2) < d_{\min}$, set $\mathbf{m} = (0,m_2)'$.
6. The current value of $\mathbf{m}$ is used to calculate $U(\mathbf{x},\mathbf{m},A,Q^+)$.

The extension of this algorithm to higher than 2 dimensions is given in the Appendix.

4. SIMULATION RESULTS AND AN EXAMPLE

To illustrate the improved power that is realized with the focused alternative hypothesis, a sequence of Monte Carlo experiments was conducted using a DEC/Alpha-2000A processor. The software was written in FORTRAN, including the pseudorandom-number generation which was transcripted from Pascal routines presented by L'Ecuyer and Cote (1991). Estimated power for each experiment was the proportion of 10,000 iterations in which the test statistic was found to be in the critical region (standard error ≤ 0.005). Various values of B, the number of bootstraps used for each focused alternative critical point, were studied. The one reported here is 499. Significance levels for the tests were $\alpha = 0.05$ and $\alpha = 0.01$. Results for $\alpha = 0.05$ are representative and are summarized in this section.

For various sample sizes from bivariate normals, Tables 1 and 2 compare the significance levels and the powers of Hotelling's T^2, Tang's half space test ($U(H_J^+)$) defined by the vector $\mathbf{J} = (1,1)'$, and our positive orthant test, $U(Q^+)$. The alternative $\boldsymbol{\mu} = (0.5,0.5)'/2^{1/2}$ is clearly in Q^+. In these tables, and in other results to follow, the diagonal elements of the covariance matrix are unity while the values of the off-diagonal elements are given by the correlation coefficient, ρ.

Table 1. Comparison of Estimated Significance Levels of Three ($\alpha = .05$) Tests in Two Dimensions with $\rho = 0$ (Standard errors are approximately .002.)

Statistic	Sample Size						
	10	17	22	27	32	42	62
$U(Q^+)$	.035	.044	.047	.051	.047	.050	.051
$U(H_J^+)$	.050	.049	.051	.052	.051	.049	.050
T^2	.052	.049	.051	.051	.050	.050	.051

Table 2a. Comparison of Estimated Powers of Three ($\alpha = .05$) Tests in Two Dimensions (Standard errors of these entries are $\leq .005$.)

($\rho = 0$)

Statistic	Sample Size						
	10	17	22	27	32	42	62
$U(Q^+)$	.252	.502	.631	.722	.795	.896	.974
$U(H_J^+)$	.263	.452	.569	.666	.737	.857	.960
T^2	.198	.372	.488	.587	.665	.803	.938

Table 2b. ($\rho = 0.5$)

Statistic	Sample Size						
	10	17	22	27	32	42	62
$U(Q^+)$	.160	.337	.448	.524	.599	.731	.883
$U(H_J^+)$	.197	.320	.420	.486	.558	.693	.858
T^2	.150	.253	.346	.407	.481	.621	.810

Table 2c. (ρ = -0.5)

Statistic	Sample Size						
	10	17	22	27	32	42	62
$U(Q^+)$	.468	.798	.904	.950	.977	.995	.999
$U(H_J^+)$	.446	.729	.852	.915	.957	.990	.999
T^2	.363	.653	.796	.880	.936	.981	.999

The observed significance levels for $\rho = -0.5$ and $+.5$ are not reported because they are similar to those in Table 1. The levels for $U(Q^+)$ are significantly less than the nominal .05 at n = 10 and 17. This suggests that the bootstrap approximation may be inadequate in the smallest samples. This conservatism accounts for it having less power than $U(H_J^+)$ at n = 10 in Tables 2a and 2b. Nevertheless it still had greater power for n 17. The same general pattern holds in Table 3.

The noncentrality parameter equals the square of the individual elements of the mean, which is 0.25 in Table 2. Maintaining this noncentrality parameter for three dimensions ($p = 3$) yields Table 3 when the correlation is zero. The standard errors are the same as in Tables 1 and 2. Because Tang did not report critical points for n = 10 and p = 3, a sample of size 11 is used in Table 3 in lieue of n = 10.

Table 3a. Comparison of Estimated Significance Levels of Three (α = .05) Tests in Three Dimensions with ρ = 0 (Standard errors are approximately .002.)

Statistic	Sample Size						
	11	17	22	27	32	42	62
$U(Q^+)$	.022	.034	.041	.044	.046	.044	.053
$U(H_J^+)$	.054	.048	.051	.051	.048	.044	.055
T^2	.055	.046	.053	.051	.049	.044	.056

Table 3b. Comparison of Estimated Powers of Three ($\alpha = .05$) Tests in Three Dimensions with $\rho = 0$ (Standard errors are $\leq .005$.)

Statistic	Sample Size						
	11	17	22	27	32	42	62
$U(Q^+)$	.162	.395	.544	.661	.746	.858	.963
$U(H_J^+)$	.211	.347	.464	.566	.648	.783	.927
T^2	.172	.292	.397	.504	.590	.734	.904

Using that same noncentrality parameter for a sample of size 27, the powers of the three tests as a function of dimension are given in Table 4. The focused test, $U(Q^+)$, has significantly greater power for every one of these cases.

Table 4. Comparison of Estimated Powers of Three ($\alpha = .05$) Tests for a Sample Size of 27 with $\rho = 0$ (Standard errors are $\leq .005$)

Statistic	Dimension				
	2	3	4	5	6
$U(Q^+)$	.722	.661	.603	.534	.476
$U(H_J^+)$	.666	.566	.490	.426	.382
T^2	.587	.504	.439	.380	.343

In a different direction, it is of interest to compare the power of $U(Q^+)$ versus $U(H_J^+)$ in the case of four dimensions and $n = 40$ for a variety of alternative means. Define the following four vectors: $\mathbf{v}_1 = (1,0,0,0)'$, $\mathbf{v}_2 = (1,1,0,0)'/2^{1/2}$, $\mathbf{v}_3 = (1,1,1,0)'/3^{1/2}$, and $\mathbf{v}_4 = (1,1,1,1)'/2$. For any of these vectors, by letting $\boldsymbol{\mu} = \lambda \mathbf{v}_i$, the power over a range of values of λ can be examined.

For negative values of λ, neither the null hypothesis nor the alternative hypothesis is true for either test. In the typical test of hypothesis situation one would wish to accept the alternative hypothesis only when it is true. Thus the contrast between the power of the two statistics presented in Table 5a is striking. This is a disturbing aspect of the test

based on $U(H_J^+)$ (see Figure 2). It can be quite likely to reject the null hypothesis even though the alternative hypothesis is far from true.

The test based on $U(Q^+)$ dramatically reduces the power outside the positive orthant, which is good. There do exist segments of the parameter space outside Q^+ where the power is high for that statistic, as well. For example, its power will increase as the true mean moves away from the point $\boldsymbol{\mu} = \mathbf{0}$ along the half space boundary $\boldsymbol{\mu}'\mathbf{J} = 0$.

When λ is positive, then the power of the half space test is effectively the same for the four $\mathbf{v}_i$, as can be seen by examining Table 5b. Note that the vector $\lambda\mathbf{v}_1$ is on the periphery of Q^+; and, as λ increases, the power of the half space statistic eventually exceeds that of the test statistic $U(Q^+)$ as the distance from the half space boundary increases. The results contained in Table 5 are also displayed in Figure 2.

Table 5a. Estimated Powers of $U(Q^+)$ and $U(H_J^+)$ when $\boldsymbol{\mu} = \lambda\mathbf{v}_i$ (Sample Size = 40)

Vector	λ = -0.8		λ = -0.6		λ = -0.4		λ = -0.2	
	$U(H_J^+)$	$U(Q^+)$	$U(H_J^+)$	$U(Q^+)$	$U(H_J^+)$	$U(Q^+)$	$U(H_J^+)$	$U(Q^+)$
$\mathbf{v}_1$	.872	.024	.623	.026	.306	.024	.096	.029
$\mathbf{v}_2$	.619	.012	.402	.012	.192	.015	.071	.016
$\mathbf{v}_3$	.283	.004	.188	.004	.093	.004	.050	.011
$\mathbf{v}_4$	.034	.000	.036	.000	.032	.000	.032	.005
	Standard error $\leq$ .005							
Observed significances (estimated standard error = .001) $U(H_J^+)$: .050 $U(Q^+)$: .045								

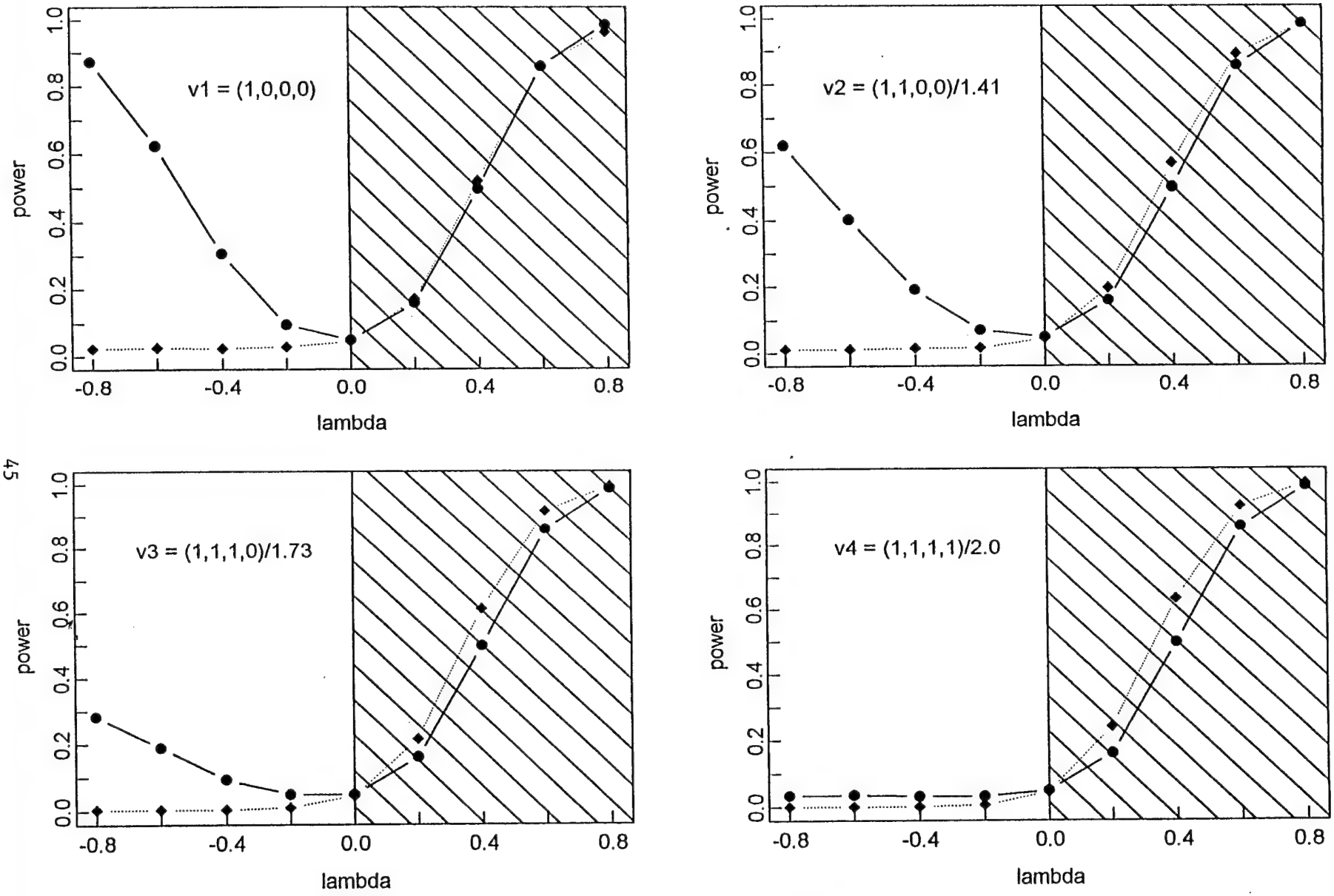

Figure 2. Estimated Power of U(Q+) (♦) and U(H+) (●) (Sample Size =40)
(Shaded Area Indicates Where Alternative Hypothesis is True)

Table 5b. Estimated Powers of $U(Q^+)$ and $U(H_J^+)$ when $\mu = \lambda \mathbf{v}_i$ (Sample Size = 40)

Vector	$\lambda = 0.2$		$\lambda = 0.4$		$\lambda = 0.6$		$\lambda = 0.8$	
	$U(H_J^+)$	$U(Q^+)$	$U(H_J^+)$	$U(Q^+)$	$U(H_J^+)$	$U(Q^+)$	$U(H_J^+)$	$U(Q^+)$
$\mathbf{v}_1$	.157	.171	.497	.520	.856	.860	.986	.963
$\mathbf{v}_2$	.164	.197	.497	.569	.856	.892	.986	.986
$\mathbf{v}_3$	.157	.215	.502	.612	.862	.915	.987	.993
$\mathbf{v}_4$	.160	.241	.497	.633	.857	.923	.986	.995
	Standard error $\leq$ .005							
Observed significances (estimated standard error = .001) $U(H_J^+)$: .050 $U(Q^+)$: .045								

EXAMPLE: In their Example 5.2, Johnson and Wichern (1988) analyze a three-dimensional data set which they call the *sweat data*. There are 20 healthy female subjects with X_1=sweat rate, X_2=sodium content, and X_3= potassium content. The hypothesized mean is μ' = (4,50,10) and by referring Hotelling's T^2 to F tables, they reject the null hypothesis at the 10% level (p=.065). To illustrate the bootstrap we suppose the alternative of interest (in advance of seeing the actual data) to be one sided and in exactly the direction of the sample mean, (4.640, 45.400, 9.965). The simple test described by Follmann (1996) yields a p-value of .0325. By recentering the raw data from their Table 5.1 and redirecting the 2nd and 3rd components, we applied our one-sided test with 30,000 bootstraps. Our estimated p-value is .028 (estimated standard error = .001), which is consistent with the notion that it is more sensitive to this class of alternatives. Again, the nonparametric bootstrap has not relied upon the multivariate normality assumption for validity, but the textbook and recent journal articles do.

5. CONCLUSION

The nonparametric bootstrap has been used to circumvent an existing serious impediment to implementing the likelihood-ratio test developed by Perlman for one-sided alternatives for the mean of a multivariate normal population with unknown covariance matrix . This was demonstrated for the specific alternative hypothesis that the mean is in the positive orthant. Empirical results verify the intuitive notion that appropriately tailoring the critical region increases the power of the test and that the relative improvement in power increases with the dimension.

Using a test statistic that has some optimal properties (at normality) as well as a very sound basis for non-normal distributions, we find exact constrained maxima by an efficient algorithm. The validity of the nonparametric resampling is remarkably good in this testing context for a range of sample sizes down to those that many would describe as small. The result is that the bootstrap can deliver greater power than the normal theory tests by being the only contender to use the relevant alternative. It is all the more noteworthy that the bootstrap did not use the fact that the data were normal and still it had greater power than the tests that used that information.

The methodology used in this paper is applicable to cones that are even more "focused" than is the positive orthant. Furthermore, the procedure is suitable for the non-normal multivariate distributions that are more realistic models in practice. The extension to the two-sample problem and the problem of detecting outliers is straightforward.

APPENDIX: Search Algorithm in p-Dimensional Space

In p-space, the process of calculating the statistic $U(\mathbf{x},\mathbf{m},A,Q^+)$ is a generalization (albeit more complicated), of what is done in the case where $p = 2$. As before, if $\mathbf{x}$ is in the positive orthant Q^+, then $\mathbf{m} = \mathbf{x}$ and $U(\mathbf{x},\mathbf{m},A,Q^+)$ is equal to the standard T^2 statistic. Otherwise, the process is to find one or more points located on the periphery of Q^+, in the highest possible dimensional subspace, which are candidates in the positive orthant, and then determine which of those points is closest to $\mathbf{x}$. This is accomplished by initially searching in $(p - 1)$ space, then $(p - 2)$ space, and so forth, until either a point is found or else it is determined that there is no positive point on any of the p axes which is closest to $\mathbf{x}$.

Denote by $\mathbf{m}_{(j)}$ a vector with the jth element being zero, i.e., $\mathbf{m}_{(j)} = (m_1,m_2,...,m_{j-1},0,m_{j+1},...,m_p)'$. The values of the elements of $\mathbf{m}_{(j)}$ are those which minimize the distance in the following equation

$$q(\mathbf{m}_{(j)}) = (\mathbf{x} - \mathbf{m}_{(j)})'A^{-1}(\mathbf{x} - \mathbf{m}_{(j)}). \qquad \text{(A.1)}$$

The elements of each $\mathbf{m}_{(j)}$ are obtained in the following manner. As before the elements of A^{-1} are denoted by c_{ij}, $i,j = 1,2,...,p$, and set

$$\mathbf{d} = A^{-1}\mathbf{x}. \qquad \text{(A.2)}$$

Let $C_{(j)}$ denote the $(p - 1)$ x $(p - 1)$ matrix obtained by deleting the jth row and column of A^{-1}, and designate by $\mathbf{d}_{(j)}$ the $(p - 1)$ vector realized when the jth element of $\mathbf{d}$ is deleted. The elements $m_1,m_2,...,m_{j-1},m_{j+1},...,m_p$ are found, respectively, as the $(p - 1)$ elements of the vector

$$\mathbf{z}_{(j)} = C^{-1}_{(ij)}\mathbf{d}_{(j)}. \qquad \text{(A.3)}$$

If any of the elements of $\mathbf{z}_{(j)}$ are negative, then the corresponding point $\mathbf{m}_{(j)}$ is not in the positive orthant. There are $\binom{p}{p-1}$ possibilities to examine. If there is more than one $\mathbf{m}_{(j)}$ in Q^+, then it can be determined which is the closest to $\mathbf{x}$.

If there is no $\mathbf{z}_{(j)}$ in which all of the elements are non-negative, then the search moves to the next lower dimension. This involves defining quantities $\mathbf{m}_{(i,j)}$, $C_{(i,j)}$, and $\mathbf{d}_{(i,j)}$, where, for example, $C_{(i,j)}$ is formed by deleting the ith and jth rows and columns of A^{-1}. Then one obtains $\mathbf{z}_{(i,j)}$:

$$\mathbf{z}_{(i,j)} = C_{(i,j)}^{-1}\mathbf{d}_{(i,j)}. \qquad \text{(A.4)}$$

There are $\binom{p}{p-2}$ possibilities to examine before moving to the next lower dimension, if required. Each time the dimension decreases, the number of elements of $\mathbf{d}$ which are used decreases by one, and one more row and column of A^{-1} are not used. Eventually, if there is no point in Q^+ which is found to be closest to $\mathbf{x}$, then $\mathbf{m} = \mathbf{0}$ and $U = 0$. This situation would require the maximum of $2^p - 1$ searches for $\mathbf{m}$.

References

Berk, R. and Marcus, R. (1996), "Dual Cones, Dual Norms, and Simultaneous Inference for Partially Ordered Means," *Journal of the American Statistical Association,* **91**, 318-328.

Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, **7**, 1 - 26.

Fisher, R. A. (1938), "The Statistical Utilization of Multiple Measurements," *Annals of Eugenics*, **8**, 376-386.

Fisher, N. I. and Hall P. (1990), "On Bootstrap Hypothesis Testing," *Australian Journal of Statistics*, **32**, 177 - 190.

Follmann, D. (1996), "A Simple Multivariate Test For One-sided Alternatives," *Journal of the American Statistical Association*, **91**, 854-861.

Johnson, R. A. and Wichern, D. W. (1988), *Applied Multivariate Statistical Analysis, 2nd Edn.*, New Jersey: Prentice Hall.

L'Ecuyer, P. and Cote, S. (1991), "Implementing a Random Number Package with Splitting Facilities," *ACM Transactions on Mathematical Software,* **17**, 98 - 111.

Perlman, M. D. (1969), "One-sided Testing Problems in Multivariate Analysis," *The Annals of Mathematical Statistics,* **40**, 549 - 567.

Tang, D. (1994), "Uniformly More Powerful Tests in a One-sided Multivariate Problem," *Journal of the American Statistical Association,* **89**, 1006 - 1011.

Titterington, D. M., Murray, G. D., Murray, L. S., Spiegelhalter, D. J., Skene, A. M., Habbema, J. D. F., and Gelpke, G. J. (1981), "Comparison of Discrimination Techniques Applied to a Complex Data Set of Head Injured Patients," *Journal of the Royal Statistical Society A*, **144**, 145 -175.

THOMAS AHRENS
SEISMOLOGICAL LABORATORY 252-21
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CA 91125

SHELTON ALEXANDER
PENNSYLVANIA STATE UNIVERSITY
DEPARTMENT OF GEOSCIENCES
537 DEIKE BUILDING
UNIVERSITY PARK, PA 16801

RICHARD BARDZELL
ACIS
DCI/ACIS
WASHINGTON, DC 20505

DOUGLAS BAUMGARDT
ENSCO INC.
5400 PORT ROYAL ROAD
SPRINGFIELD, VA 22151

WILLIAM BENSON
NAS/COS
ROOM HA372
2001 WISCONSIN AVE. NW
WASHINGTON, DC 20007

ROBERT BLANDFORD
AFTAC
1300 N. 17TH STREET
SUITE 1450
ARLINGTON, VA 22209-2308

RHETT BUTLER
IRIS
1200 NEW YORK AVE., NW
SUITE 800
WASHINGTON, DC 20005

CATHERINE DE GROOT-HEDLIN
UNIVERSITY OF CALIFORNIA, SAN DIEGO
INSTITUTE OF GEOPHYSICS AND PLANETARY PHYSICS
8604 LA JOLLA SHORES DRIVE
SAN DIEGO, CA 92093

SEAN DORAN
ACIS
DCI/ACIS
WASHINGTON , DC 20505

RICHARD J. FANTEL
BUREAU OF MINES
DEPT OF INTERIOR, BLDG 20
DENVER FEDERAL CENTER
DENVER, CO 80225

RALPH ALEWINE
NTPO
1901 N. MOORE STREET, SUITE 609
ARLINGTON, VA 22209

MUAWIA BARAZANGI
INSTITUTE FOR THE STUDY OF THE CONTINENTS
3126 SNEE HALL
CORNELL UNIVERSITY
ITHACA, NY 14853

T.G. BARKER
MAXWELL TECHNOLOGIES
P.O. BOX 23558
SAN DIEGO, CA 92123

THERON J. BENNETT
MAXWELL TECHNOLOGIES
11800 SUNRISE VALLEY DRIVE SUITE 1212
RESTON, VA 22091

JONATHAN BERGER
UNIVERSITY OF CA, SAN DIEGO
SCRIPPS INSTITUTION OF OCEANOGRAPHY IGPP, 0225
9500 GILMAN DRIVE
LA JOLLA, CA 92093-0225

STEVEN BRATT
NTPO
1901 N. MOORE STREET, SUITE 609
ARLINGTON, VA 22209

LESLIE A. CASEY
DOE
1000 INDEPENDENCE AVE. SW
NN-20
WASHINGTON, DC 20585-0420

STANLEY DICKINSON
AFOSR
110 DUNCAN AVENUE, SUITE B115
BOLLING AFB
WASHINGTON, D.C. 20332-001

DIANE I. DOSER
DEPARTMENT OF GEOLOGICAL SCIENCES
THE UNIVERSITY OF TEXAS AT EL PASO
EL PASO, TX 79968

JOHN FILSON
ACIS/TMG/NTT
ROOM 6T11 NHB
WASHINGTON, DC 20505

MARK D. FISK
MISSION RESEARCH CORPORATION
735 STATE STREET
P.O. DRAWER 719
SANTA BARBARA, CA 93102-0719

LORI GRANT
MULTIMAX, INC.
311C FOREST AVE. SUITE 3
PACIFIC GROVE, CA 93950

I. N. GUPTA
MULTIMAX, INC.
1441 MCCORMICK DRIVE
LARGO, MD 20774

IAN MACGREGOR
NSF
4201 WILSON BLVD., ROOM 785
ARLINGTON, VA 22230

MICHAEL HEDLIN
UNIVERSITY OF CALIFORNIA, SAN DIEGO
SCRIPPS INSTITUTION OF OCEANOGRAPHY IGPP, 0225
9500 GILMAN DRIVE
LA JOLLA, CA 92093-0225

EUGENE HERRIN
SOUTHERN METHODIST UNIVERSITY
DEPARTMENT OF GEOLOGICAL SCIENCES
DALLAS, TX 75275-0395

VINDELL HSU
HQ/AFTAC/TTR
1030 S. HIGHWAY A1A
PATRICK AFB, FL 32925-3002

RONG-SONG JIH
PHILLIPS LABORATORY
EARTH SCIENCES DIVISION
29 RANDOLPH ROAD
HANSCOM AFB, MA 01731-3010

LAWRENCE LIVERMORE NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 808, MS L-200
LIVERMORE, CA 94551

LAWRENCE LIVERMORE NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
LLNL
PO BOX 808, MS L-175
LIVERMORE, CA 94551

ROBERT GEIL
DOE
PALAIS DES NATIONS, RM D615
GENEVA 10, SWITZERLAND

HENRY GRAY
SMU STATISTICS DEPARTMENT
P.O. BOX 750302
DALLAS, TX 75275-0302

DAVID HARKRIDER
PHILLIPS LABORATORY
EARTH SCIENCES DIVISION
29 RANDOLPH ROAD
HANSCOM AFB, MA 01731-3010

THOMAS HEARN
NEW MEXICO STATE UNIVERSITY
DEPARTMENT OF PHYSICS
LAS CRUCES, NM 88003

DONALD HELMBERGER
CALIFORNIA INSTITUTE OF TECHNOLOGY
DIVISION OF GEOLOGICAL & PLANETARY SCIENCES
SEISMOLOGICAL LABORATORY
PASADENA, CA 91125

ROBERT HERRMANN
ST. LOUIS UNIVERSITY
DEPARTMENT OF EARTH & ATMOSPHERIC SCIENCES
3507 LACLEDE AVENUE
ST. LOUIS, MO 63103

ANTHONY IANNACCHIONE
BUREAU OF MINES
COCHRANE MILL ROAD
PO BOX 18070
PITTSBURGH, PA 15236-9986

THOMAS JORDAN
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
EARTH, ATMOSPHERIC & PLANETARY SCIENCES
77 MASSACHUSETTS AVENUE, 54-918
CAMBRIDGE, MA 02139

LAWRENCE LIVERMORE NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 808, MS L-221
LIVERMORE, CA 94551

LAWRENCE LIVERMORE NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 808, MS L-208
LIVERMORE, CA 94551

LAWRENCE LIVERMORE NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 808, MS L-202
LIVERMORE, CA 94551

LAWRENCE LIVERMORE NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 808, MS L-205
LIVERMORE, CA 94551

ANATOLI L. LEVSHIN
DEPARTMENT OF PHYSICS
UNIVERSITY OF COLORADO
CAMPUS BOX 390
BOULDER, CO 80309-0309

LOS ALAMOS NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 1663, MS F659
LOS ALAMOS, NM 87545

LOS ALAMOS NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 1663, MS D460
LOS ALAMOS, NM 87545

GARY MCCARTOR
SOUTHERN METHODIST UNIVERSITY
DEPARTMENT OF PHYSICS
DALLAS, TX 75275-0395

BRIAN MITCHELL
DEPARTMENT OF EARTH & ATMOSPHERIC SCIENCES
ST. LOUIS UNIVERSITY
3507 LACLEDE AVENUE
ST. LOUIS, MO 63103

JOHN MURPHY
MAXWELL TECHNOLOGIES
11800 SUNRISE VALLEY DRIVE SUITE 1212
RESTON, VA 22091

JOHN ORCUTT
INSTITUTE OF GEOPHYSICS AND PLANETARY PHYSICS
UNIVERSITY OF CALIFORNIA, SAN DIEGO
LA JOLLA, CA 92093

PACIFIC NORTHWEST NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 999, MS K7-34
RICHLAND, WA 99352

LAWRENCE LIVERMORE NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 808, MS L-195
LIVERMORE, CA 94551

THORNE LAY
UNIVERSITY OF CALIFORNIA, SANTA CRUZ
EARTH SCIENCES DEPARTMENT
EARTH & MARINE SCIENCE BUILDING
SANTA CRUZ, CA 95064

DONALD A. LINGER
DNA
6801 TELEGRAPH ROAD
ALEXANDRIA, VA 22310

LOS ALAMOS NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 1663, MS F665
LOS ALAMOS, NM 87545

LOS ALAMOS NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 1663, MS C335
LOS ALAMOS, NM 87545

KEITH MCLAUGHLIN
MAXWELL TECHNOLOGIES
P.O. BOX 23558
SAN DIEGO, CA 92123

RICHARD MORROW
USACDA/IVI
320 21ST STREET, N.W.
WASHINGTON, DC 20451

JAMES NI
NEW MEXICO STATE UNIVERSITY
DEPARTMENT OF PHYSICS
LAS CRUCES, NM 88003

PACIFIC NORTHWEST NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 999, MS K6-48
RICHLAND, WA 99352

PACIFIC NORTHWEST NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 999, MS K6-40
RICHLAND, WA 99352

PACIFIC NORTHWEST NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 999, MS K6-84
RICHLAND, WA 99352

FRANK PILOTTE
HQ/AFTAC/TT
1030 S. HIGHWAY A1A
PATRICK AFB, FL 32925-3002

JAY PULLI
BBN
1300 NORTH 17TH STREET
ROSSLYN, VA 22209

DAVID RUSSELL
HQ AFTAC/TTR
1030 SOUTH HIGHWAY A1A
PATRICK AFB, FL 32925-3002

SANDIA NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
DEPT. 5704
MS 0979, PO BOX 5800
ALBUQUERQUE, NM 87185-0979

SANDIA NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
DEPT. 9311
MS 1159, PO BOX 5800
ALBUQUERQUE, NM 87185-1159

SANDIA NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
DEPT. 5736
MS 0655, PO BOX 5800
ALBUQUERQUE, NM 87185-0655

AVI SHAPIRA
SEISMOLOGY DIVISION
THE INSTITUTE FOR PETROLEUM RESEARCH AND GEOPHYSICS
P.O.B. 2286, NOLON 58122 ISRAEL

MATTHEW SIBOL
ENSCO, INC.
445 PINEDA COURT
MELBOURNE, FL 32940

JEFFRY STEVENS
MAXWELL TECHNOLOGIES
P.O. BOX 23558
SAN DIEGO, CA 92123

PACIFIC NORTHWEST NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
PO BOX 999, MS K5-12
RICHLAND, WA 99352

KEITH PRIESTLEY
DEPARTMENT OF EARTH SCIENCES
UNIVERSITY OF CAMBRIDGE
MADINGLEY RISE, MADINGLEY ROAD
CAMBRIDGE, CB3 OEZ UK

PAUL RICHARDS
COLUMBIA UNIVERSITY
LAMONT-DOHERTY EARTH OBSERVATORY
PALISADES, NY 10964

CHANDAN SAIKIA
WOOODWARD-CLYDE FEDERAL SERVICES
566 EL DORADO ST., SUITE 100
PASADENA, CA 91101-2560

SANDIA NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
DEPT. 5791
MS 0567, PO BOX 5800
ALBUQUERQUE, NM 87185-0567

SANDIA NATIONAL LABORATORY
ATTN: TECHNICAL STAFF (PLS ROUTE)
DEPT. 5704
MS 0655, PO BOX 5800
ALBUQUERQUE, NM 87185-0655

THOMAS SERENO JR.
SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
10260 CAMPUS POINT DRIVE
SAN DIEGO, CA 92121

ROBERT SHUMWAY
410 MRAK HALL
DIVISION OF STATISTICS
UNIVERSITY OF CALIFORNIA
DAVIS, CA 95616-8671

DAVID SIMPSON
IRIS
1200 NEW YORK AVE., NW
SUITE 800
WASHINGTON, DC 20005

BRIAN SULLIVAN
BOSTON COLLEGE
INSITUTE FOR SPACE RESEARCH
140 COMMONWEALTH AVENUE
CHESTNUT HILL, MA 02167

DAVID THOMAS
ISEE
29100 AURORA ROAD
CLEVELAND, OH 44139

LAWRENCE TURNBULL
ACIS
DCI/ACIS
WASHINGTON, DC 20505

FRANK VERNON
UNIVERSITY OF CALIFORNIA, SAN DIEGO
SCRIPPS INSTITUTION OF OCEANOGRAPHY IGPP, 0225
9500 GILMAN DRIVE
LA JOLLA, CA 92093-0225

DANIEL WEILL
NSF
EAR-785
4201 WILSON BLVD., ROOM 785
ARLINGTON, VA 22230

RU SHAN WU
UNIVERSITY OF CALIFORNIA SANTA CRUZ
EARTH SCIENCES DEPT.
1156 HIGH STREET
SANTA CRUZ, CA 95064

JAMES E. ZOLLWEG
BOISE STATE UNIVERSITY
GEOSCIENCES DEPT.
1910 UNIVERSITY DRIVE
BOISE, ID 83725

DEFENSE TECHNICAL INFORMATION CENTER
8725 JOHN J. KINGMAN ROAD
FT BELVOIR, VA 22060-6218 (2 COPIES)

PHILLIPS LABORATORY
ATTN: GPBP
29 RANDOLPH ROAD
HANSCOM AFB, MA 01731-3010

PHILLIPS LABORATORY
ATTN: RESEARCH LIBRARY/TL
5 WRIGHT STREET
HANSCOM AFB, MA 01731-3004

NAFI TOKSOZ
EARTH RESOURCES LABORATORY, M.I.T.
42 CARLTON STREET, E34-440
CAMBRIDGE, MA 02142

GREG VAN DER VINK
IRIS
1200 NEW YORK AVE., NW
SUITE 800
WASHINGTON, DC 20005

TERRY WALLACE
UNIVERSITY OF ARIZONA
DEPARTMENT OF GEOSCIENCES
BUILDING #77
TUCSON, AZ 85721

JAMES WHITCOMB
NSF
NSF/ISC OPERATIONS/EAR-785
4201 WILSON BLVD., ROOM785
ARLINGTON, VA 22230

JIAKANG XIE
COLUMBIA UNIVERSITY
LAMONT DOHERTY EARTH OBSERVATORY
ROUTE 9W
PALISADES, NY 10964

OFFICE OF THE SECRETARY OF DEFENSE
DDR&E
WASHINGTON, DC 20330

TACTEC
BATTELLE MEMORIAL INSTITUTE
505 KING AVENUE
COLUMBUS, OH 43201 (FINAL REPORT)

PHILLIPS LABORATORY
ATTN: GPE
29 RANDOLPH ROAD
HANSCOM AFB, MA 01731-3010

PHILLIPS LABORATORY
ATTN: PL/SUL
3550 ABERDEEN AVE SE
KIRTLAND, NM 87117-5776 (2 COPIES)